Analysis of Washington Nutrient and Biological Data (Periphyton) for the Nutrient Scientific Technical Exchange Partnership Support (N-STEPS)

Prepared for:

U.S. Environmental Protection Agency
Office of Science and Technology,
Health Ecological Criteria Division
1200 Pennsylvania Avenue, NW
Washington, DC 20460

And

Office of Water and Watersheds
U.S. Environmental Protection Agency, Region 10
1200 Sixth Avenue, Suite 900
Seattle, WA 98101-3140

Prepared by:

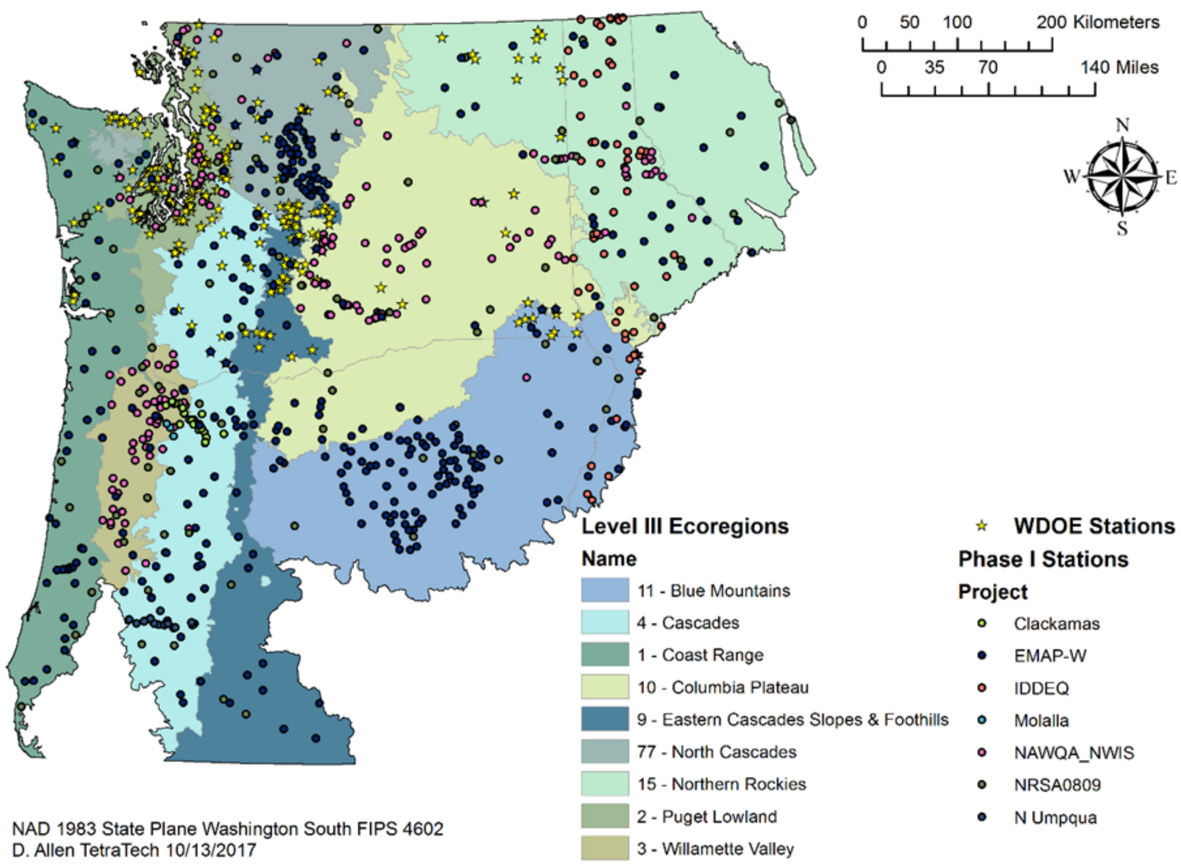
Tetra Tech, Inc., Center for Ecological Sciences
1 Park Drive, Suite 200
Research Triangle Park, NC 27709

April 19, 2018

Contents

Executive Summary

This report summarizes the results of an analysis of water chemistry and stream periphyton data for the State of Washington. The goals of this research were to: 1) evaluate if the periphyton assemblage could be used as a screening tool to identify nutrient conditions in Washington streams, and 2) explore numeric nutrient thresholds associated with periphyton responses to nutrient gradients in streams. This analysis built upon previous work for the state of Oregon, for which western regional EPA and USGS paired nutrient and periphyton data were compiled and used to address identical questions in Oregon. For this effort, Washington Department of Ecology (ECY) paired nutrient and periphyton data were added to the prior effort dataset and the combined data reanalyzed. **The resulting dataset had more than 4500 samples from more than 850 sites, including nearly 1000 paired nutrient-diatom samples.** As a result, insights from these analyses are robust. Additional data may provide further insight, but **these sample sizes are sufficient for making inferences from the resulting relationships and conclusions and comprise all high quality available data that could be identified**.



Sample locations used in this analysis

We conducted exploratory analysis to characterize nutrient distributions across the state. In general, **higher nutrient concentrations and nutrient imbalances were observed in the**

Puget Lowlands and Columbia Plateau and lower nutrient concentrations were observed in the Cascades. Median total nitrogen concentrations were highest in the Puget Lowlands, followed by the Columbia Plateau, Blue Mountains, and Coast Range. Lowest total nitrogen concentrations were found in the Cascades ecoregion. Median total phosphorus concentrations were highest in the Puget Lowlands, Columbia Plateau, and Willamette Valley and lowest total phosphorus concentrations were observed in the Cascades. Nutrient imbalances, like those observed in the Puget Lowlands, can be associated with changes in algal assemblages and ecosystem processes.

Because natural variability due to biogeographic gradients in nutrient or nutrient-response relationships introduces controllable error, classification analyses were conducted to see if the state of Washington could be split into classes within which nutrient dynamics and periphyton responses to nutrient gradients are naturally different. The result of this analysis was to preliminarily **recommend that the Willamette Valley, Columbia Plateau, and Blue Mountains ecoregions (ecoregions 3, 10, and 11 respectively) be considered separately for developing screening tools and numeric thresholds for nitrogen but not necessarily for phosphorus.** Phosphorus models were similar enough among different classes to be combined.

We then conducted stressor-response relationship modeling using linear regression to model the response of several diatom metrics to nutrient gradients. The diatom metrics we used were developed independently and are based on diatom taxa sensitivities to factors like land disturbance, embeddedness, conductivity, and nutrient enrichment. Existing **diatom metrics were significantly sensitive indicators of both phosphorus and nitrogen pollution**, although more so to phosphorus. Nutrients explained anywhere from 25 to 50% of the variability in diatom metrics, which is generally quite high for stream biological responses. In a separate analysis for the Oregon project, we also discovered these responses are robust to covariates, meaning the variability explained by nutrients are not confounded, but represent true contributions. This means these **diatom metrics are appropriate and defensible screening tools for nutrient pollution**.

To inform the selection of candidate nutrient thresholds, we took a reference site benchmark approach. We identified least disturbed reference sites based on program-specific reference site designations. We then calculated quartile and decile reference site diatom metric values as benchmarks. The target values represent the point beyond which a sample is no longer like the reference population. We then interpolated from the regression models, the nutrient concentrations associated with this biological target and **identified nutrient concentrations where stream diatoms were no longer like reference sites**. These values could be considered for use as screening or management targets for protection of aquatic life conditions in a stream.

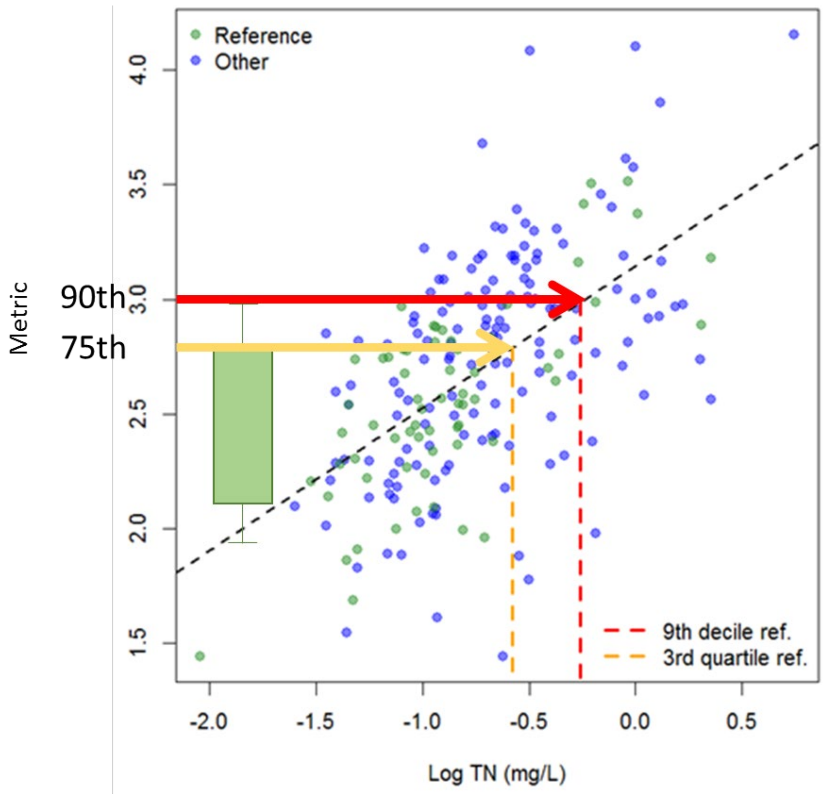


Figure demonstrating conceptual approach for deriving numeric nutrient values associated with reference based diatom metric benchmarks.

Numeric nutrient values as low as 0.01 mg/L (range: 0.010 to 0.080) total phosphorus and 0.060 (range: 0.060 to 0.410) mg/L total nitrogen were associated with the 75th percentile of reference metric scores. Such values are low, but within the range for total phosphorus reported from other published studies of algal metric responses to nutrients, while on the low end of those reported for total nitrogen.

Among aquatic life, diatoms are the most immediate responders to nutrient enrichment in streams. Changes in diatom assemblage structure are expected to presage further changes including proliferation of nuisance or harmful species. Diatoms also represent a significant source of energy for stream food webs and changes in diatom assemblages can have bottom-up effects on stream communities. As a result, changes in diatom assemblage structure also presage changes in macroinvertebrate or fish assemblages. Finally, because changes in diatom structure precede increases in algal biomass and the metabolic changes that increased biomass can produce (i.e., high and low DO), these diatom structural changes also presage changes in DO and pH dynamics. For these reasons, **diatoms are important early indicators of stress from nutrient pollution**. Moreover, EPA's critical elements review program gives highest assessment program scores to those incorporating all 3 main assemblages in streams: algae, invertebrates, and fish. In combination with the above sensitivity to nutrients, **algae are an important biological indicator for ECY's monitoring and assessment program**.

NSTEPS believes the following recommendations merit consideration:

i. Incorporate diatoms and other algae as part of routine monitoring.
ii. Use these data to support decisions related to nutrients and other impacts as you would invertebrate or fish information.
iii. Consider the reference based numeric diatom metric benchmark values and the interpolated numeric nutrient thresholds derived from this analysis as screening levels for indicating incipient impacts to biota.

NSTEPS recommends ECY consider the following as next steps to further improve this analysis:

i. Resolve issues of different reference condition in the Puget Lowlands resulting from different screening requirements from other ecoregions required to have sufficient reference sites in this region; reconcile differences in expectations resulting from the different screening requirements.
ii. Decide on statewide total phosphorus versus distinct total nitrogen classification and whether having classes for nitrogen and combined approach for phosphorus is acceptable. This may involve continued exploration using additional classification analyses.
iii. Given that changes in diatom assemblage presage other ecosystem changes, consider confirming and quantifying the relationship between diatom metrics and other aquatic life use measures including invertebrate and fish measures, dissolved oxygen, biomass, etc. to further tie the diatom metric benchmarks to other desired aquatic life use assessment endpoints for WA streams.
iv. Consider additional validation by testing the relationships with a new or independent dataset. However, NSTEPS believes the size of this dataset and the validation of these models with independent data from Oregon confirm that these existing models are sufficiently robust.
v. Consider conducting nutrient inferential modeling. Diatom taxa have nutrient optima and can be used to infer the actual average nutrient conditions at sites, which may contrast distinctly from measured nutrient grab samples. New Jersey uses this approach as part of the diatom indicator they developed and use in screening for nutrient impacts in streams.

1. INTRODUCTION

This report summarizes results of the workplan developed to analyze existing water chemistry and periphyton data from the State of Washington. The main goal of this research was to evaluate how well the periphyton assemblage could be used as a screening tool to identify excess nutrient conditions in Washington streams. In addition, because natural variability due to biogeographic gradients in nutrient or nutrient-response relationships introduces controllable error, the effects of a priori classes (e.g., ecoregions) and available natural gradient factors were explored for potential division of Washington sites into separate classes. If classes are identified, classification will reduce variability in distributional statistics and improve fit of stressor-response models. This work builds from an effort to look at similar responsiveness of periphyton metrics to nutrient gradients in Oregon, incorporating additional data from the State of Washington, Department of Ecology.

The proposed outcome of this analysis was to develop a brief report (this document) that provided the following:

1. A conceptual model for stressor-response relationships for the stream types (subclasses) that ECY identifies (Appendix 1).
2. A GIS map of relevant locations with applicable data throughout the western ecoregions.
3. Site classes and co-varying environmental variables used to reduce natural variability in the nutrient data developed from the broader (multi-state) dataset.
4. Analysis of the above biological samples (periphyton identification and quantification), including:
    - Endpoints determined from frequency distribution analysis for TN and TP, including cumulative distribution functions and the range of standard distributional statistics (mean, variance, standard deviation, standard errors, coefficient of variation, median, and quartiles);
    - Endpoints from stressor response analysis including visual plots of interest with loess curve fits, interpolated linear regression stressor values for response of interest, and thresholds determined for non-linear relationships using change-point analysis. Uncertainty statistics for each estimate;
    - A qualitative discussion of data gaps and the basis for any recommendations. Gaps discussion will focus on ranges of gradients and physical locations, but will also include analysis gaps that could be addressed with future data, modeling, or field collection efforts;
5. A database containing relevant data that is gleaned from reported studies/assessments, if requested.
6. All R code, annotated with detail, for analyses will also be provided, if requested.

Item 1 was provided as a series of powerpoint presentations on the conceptual model (see Appendix 1). This report represents provision of items 2-4, and items 5 and 6 will be provided or made available along with this report as requested.

The report briefly discusses methods, results, and provides a summary of the data analysis. Example plots and tables are shown in the main text, additional data and analysis output are provided in the appendices.

2. METHODS

Study area

The initial study area included the state of Washington and surrounding sampling stations that are located within Omernik level III ecoregions (USEPA, 2013) that are also contained within the state of Washington (Figure 1). After preliminary classification analysis, however, in which latitude was shown to have an influence on diatom metric vs. nutrient concentration model residuals and concern about extensive distance creating non-relevant effects on community structure, California sites (latitude < 42°N) were removed from subsequent analyses and analyses were repeated without those sites.

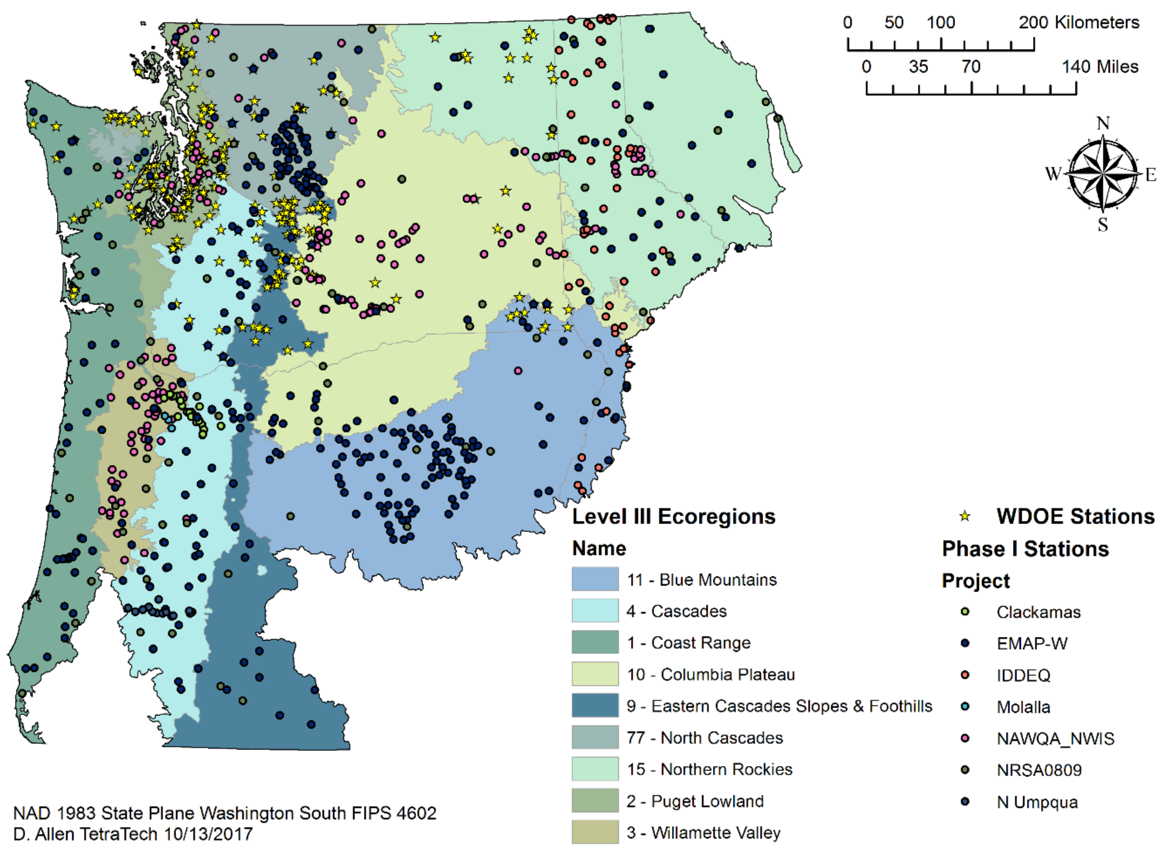


Figure 1 - Sampling stations (N=856) included in the current analysis mapped along with Washington ecoregions.

Data preparation

Stream chemistry, nutrient, and periphyton data were obtained from two sources: 1) Washington Department of Ecology (ECY) and 2) a previous NSTEPS project developed for the state of Oregon which contained data from several sources (Table 1). The ECY data were merged with the previous NSTEPS project data. Water quality parameters used in this analysis were limited to those provided by these sources (Table 2). Because EPA and USGS projects used different nomenclature, chemical parameters collected in those projects were aggregated (Appendix 2). Some values derived from older analytical techniques were excluded as the values are not comparable to newer methods. Samples collected earlier than 2000 were also excluded to better match the ECY dataset, whose samples were collected from 2010 - 2015. Samples whose shipping or holding times exceeded recommended ranges were included. Because reported values were frequently below detection limits, and detection limits varied by project and within parameters, all values were included as recorded in the original datasets. For the ECY data, the values were used as provided without modification.

All chemical parameters except pH, dissolved oxygen (DO), and temperature were log transformed. Averages therefore often represent geometric means. The final dataset prepared by the methods described above included 4668 samples from 856 distinct sampling sites.

Table 1 – Source of data used in the analysis.

Agency	Project	Data Source
Washington Department of Ecology	Multiple	ECY, Chad Larson
US EPA	Western EMAP (Environmental Monitoring and Assessment Program)	http://www.epa.gov/wed/pages/models/EMAP_West_Data.htm
	NRSA0809 (National Rivers and Streams Assessment)	http://water.epa.gov/type/rsl/monitoring/riverssurvey/index.cfm
USGS Special Projects	Clackamas	USGS, Kurt Carpenter
	Molalla	USGS, Kurt Carpenter
	N. Umpqua	USGS, Kurt Carpenter
USGS	NAWQA (National Water Quality Assessment)	USGS, Kurt Carpenter
Idaho Department of Environmental Quality		Tetra Tech, Ben Jessup

For the periphyton data collected by non-ECY entities, two or three samples had frequently been collected on each visit to a given site. To avoid duplication, only one sample per site visit was chosen to include in the analysis. For those samples for which the sample type was specified, types ID and ARTH were included. For samples collected at the same site on the same day, samples labelled "primary" or "A" were included and samples labelled "duplicate" or "B" were excluded. Diatom counts including less than 500 or more than 700 total valves counted were excluded. Several samples included duplicate species counts; those counts were summed.

Additional diatom sample preparation unique to the ECY samples included: adding a code for diatom only samples, removing all NA taxa from consideration, fixing SampleID - QUAR77-2015-0715-10:11 so as to be included, and merging taxa that had two separate enumerations within a sample for the same taxon.

After all taxonomic data from ECY were merged with the previous data, a taxonomic reconciliation effort was undertaken. Because samples were obtained from several labs, nomenclature varied or ambiguous taxonomies were often used (e.g., resolution of some individuals to genus and others to species within a genus, etc.). Therefore, these differences had to be resolved and the taxonomy converted to the nomenclature tied to the diatom attribute and trait information used to calculate metrics.

Periphyton data were collected into one table. In general, reported taxa containing "cf.," "aff.," or "sp." with or without a number or question mark, were considered ambiguous identifications and were assigned the lowest taxonomic identification possible (usually genus). Reported taxa that were unambiguously identified were compared to the Academy of Natural Sciences [ANS; NADED (North American Diatom Ecological Database) fields] taxa lists and to taxonomic names previously reviewed by Tetra Tech. If a taxon name appeared on the ANS list, the taxon was assigned that name. If not, but it did appear on the Tetra Tech list, it was assigned that name. If it appeared on neither list, it was compared to the California Academy of Sciences (CAAS) database, and if a match was found, was assigned that taxonomic name. Finally, remaining taxa were searched for on ITIS (www.itis.gov) and AlgaeBase (www.algaebase.org) to identify possible synonyms or basionyms that might be present on one of the above lists. If no matches were found, the taxon was resolved to genus if possible or excluded if not. After taxon reconciliation above, some taxa were listed twice, ~~for~~ example, two separate taxa resolved up the same genus. Those counts were combined, and then total counts and relative abundances were recalculated for all the samples. In the end, 46 taxa remained unresolved and were removed, these were rare taxa however and represented 0.3% of individual valves and affected only 0.5% of the samples.

From the diatom data, 30 metrics were calculated using an Access database developed for a prior project (Appendix 3). These primarily consisted of two main metric types: composition ("percent individuals") and weighted average based. The composition metrics are calculated as the percent of individual diatoms in the sample reflecting a particular trait, for example benthic or sestonic algae for the BEN_SES metric. The weighted average based metrics are ones that are calculated as a weighted average of taxa optima (using relative abundance) to specific nutrient or stressor (e.g., total phosphorus, TP) gradients. An example is the wa_OptCat_L1Ptl which is a diatom metric that is a relative abundance weighted TP optima index based on individual diatom species TP optima. A second is the wa_OptCat_NutMMI which is a diatom metric that is a relative abundance weighted multivariate nutrient optima index based on individual diatom species optima to a combined TP and TN gradient. Weighted average metrics were also calculated for some specific trait categories such as MOISTURE, pH, POLL_TOL where the result is s weighted average index value based on diatom relative abundances and species specific trait values for desiccation, pH, and general pollution tolerance, respectively.

To match periphyton data to water chemistry samples, a window of 28 days before and 14 days after was used to choose candidate chemistry samples. From that window, the sample closest in time to the periphyton sample collection date was manually selected. For samples without a collection date (USGS special projects), samples were matched manually by year. The dataset described above containing the periphyton assemblage metrics and paired chemistry data contained 880 samples representing 674 distinct sampling sites.

Lastly, for classification analysis, physical, geographic and landscape-level variables were collected. Precipitation was extracted from the PRISM Climate Group 30 year (1981-2010) normal annual precipitation raster (2012). For landscape level characteristics (i.e. forest cover), ECY provided StreamCat data, which provides summary statistics based on the watershed area of each NHD reach on which a sampling station is located. For the non-ECY data, watersheds had been previously delineated to calculate landcover characteristics based on the watershed upstream of a given sampling sites. For example, for sampling stations located near the downstream node of an NHD reach, the delineated watershed would provide nearly identical results to the data provided by StreamCat for that reach. For a sampling site further upstream from the outlet, the delineated watershed would not include contributing watershed area to the reach downstream of the site in the calculated land cover data.

Data description, visualization and frequency distribution

Descriptive statistics were calculated for each parameter. Boxplots and cumulative distribution functions were created for all parameters for all the study sites by reference designation, and for total phosphorus (TP) and total nitrogen (TN) for study sites by ecoregion. Reference site designations were used as attributed by ECY, EMAP-W and NRSA in their datasets. However, because some projects from the previous NSTEPS database did not include a reference designation, some of the “Other” category streams may be from relatively undisturbed watersheds. Samples from the Omernik Level III Ecoregion 2 (Puget Lowland) were plotted separately as their reference sites had higher TN concentrations than reference sites from other ecoregions and ECY scientists indicated different reference screening values were used for ecoregion 2 in order to have a sufficient reference site population.

Stressor-Response

Spearman correlation analysis was used to explore relationships between biological response periphyton metrics and nutrient concentrations. Loess (locally estimated scatterplot smoothing) fits were used to explore the shape of relationships. Then, linear models were developed to determine stressor-response relationships. Where statistically significant ($p < 0.05$) nutrient-periphyton relationships existed, the nutrient concentrations associated with response targets calculated from reference sites (10^{th} and 25^{th} percentiles for linear models with negative slopes and 75^{th} and 90^{th} percentiles for linear models with positive slopes) were interpolated. Significant correlations ($p>0.05$), were selected from among the periphyton metric response models and used for guiding subsequent classification analyses.

Following classification analysis, stressor response analyses were repeated for the candidate classes.

Classification

Natural variability due to biogeographic gradients in nutrient or nutrient-response relationships introduces controllable error. Therefore, the effect of a priori classes (e.g., ecoregions) and available natural gradient factors were explored for potential classifications to reduce variability in distributional statistics or stressor-response models. Three general techniques were used to assess the effect of ecoregions and natural gradients on model relationships as separate and supporting lines of evidence: 1) visual analysis of model residuals, 2) model-based recursive partitioning, and 3) Nash-Sutcliffe efficiency calculation. Natural gradients included elevation, precipitation, longitude, latitude and percent forest cover.

To examine model residuals in relation to environmental gradients, the periphyton metrics with the highest Spearmen correlations to nutrient concentrations were chosen to develop test models. Those models were run with all the data, and model residuals were examined as a function of the following predictors: ecoregion, longitude, latitude, elevation, and precipitation. Visual trends in residuals with any predictor indicate a potential effect of natural gradients on differences in nutrient response.

Model-based recursive partitioning was used to determine if and how biogeographic gradients and ecoregion classes affect the linear regression model between nutrient concentrations and periphyton metrics (Alexander and Grimshaw, 1996). Model-based recursive partitioning splits samples into subsets, based on a specified continuous or categorical covariate, in which the relationship between the stressor and response variables has minimal summed deviance across the individual models for all subsets of the data as compared to the deviance of the original model built from unsplit data (Zeileis et al 2008). To implement model-based partitioning, the partykit package (Hothorn and Zeileis 2014) in R was used (R Core Team 2014).

After determining possible classes, Nash-Sutcliffe efficiency (NSE) was calculated to quantify how well diatom metric - nutrient concentration models built from one class described relationships from sites of a different class. The NSE is a normalized statistic that determines the relative magnitude of a test group residual variance compared to the measured data variance with a “1” being a perfect match of modelled to observed data, and values less than one indicating that the data mean is a better predictor than the modelled values (Nash and Sutcliffe 1970). It is calculated by:

NSE = 1 - (sum((obs - sim)^2) / sum((obs - mean(obs))^2)

NSE was calculated for diatom metric vs. nutrient concentration models for ecoregions 3 (Willamette Valley), 10 (Columbia Plateau), and 11 (Blue Mountains) combined, to which samples from remaining ecoregions were compared and NSE of that test data compared to the ecoregion 3, 10, and 11 model. The hydroGOF package was used to calculate NSE (Zambrano-Bigiarini 2014).

3. RESULTS

Data description, visualization and frequency distribution

Exploratory data visualization revealed that reference sites in the Puget Lowland (ecoregion 2) had higher TN values than other ecoregions, although values across all sites did not vary as much (Figure 2). ECY confirmed that due to widespread urbanization in the Puget Lowland ecoregion, it was necessary to relax reference site screening values to some degree in this ecoregion to provide a sufficient reference site population. Therefore, in summarizing nutrient concentrations by reference and non-reference populations, ecoregion 2 reference sites were evaluated separately. Ecoregions 1 (Coast Range), 10 (Columbia Plateau) and 11 (Blue Mountains) tended to have higher TN concentrations in reference sites, whereas TP values in reference sites were higher at ecoregion 9 (Eastern Cascades Slopes and Foothills), 10 (Columbia Plateau), and 11 (Blue Mountains) sites (Figure 2 and Figure 3). Ecoregion specific cumulative frequency distributions across all sites (including non-reference) were otherwise characterized by higher nutrients in ecoregions 3 and 10 (Willamette Valley and Columbia Plateau).

Descriptive statistics of grab samples for each parameter from all sites, reference sites and reference sites in ecoregion 2 sites (Puget Lowland) are summarized in Table 2 and graphically in Figure 4 to Figure 7. As expected, nutrient concentrations are lower in reference sites (not including ecoregion 2). Ecoregion 2 (Puget Lowland) reference sites had higher mean chloride (Cl), total suspended solids (TSS), specific conductivity, turbidity, and benthic chlorophyll *a* biomass than sites not designated reference.

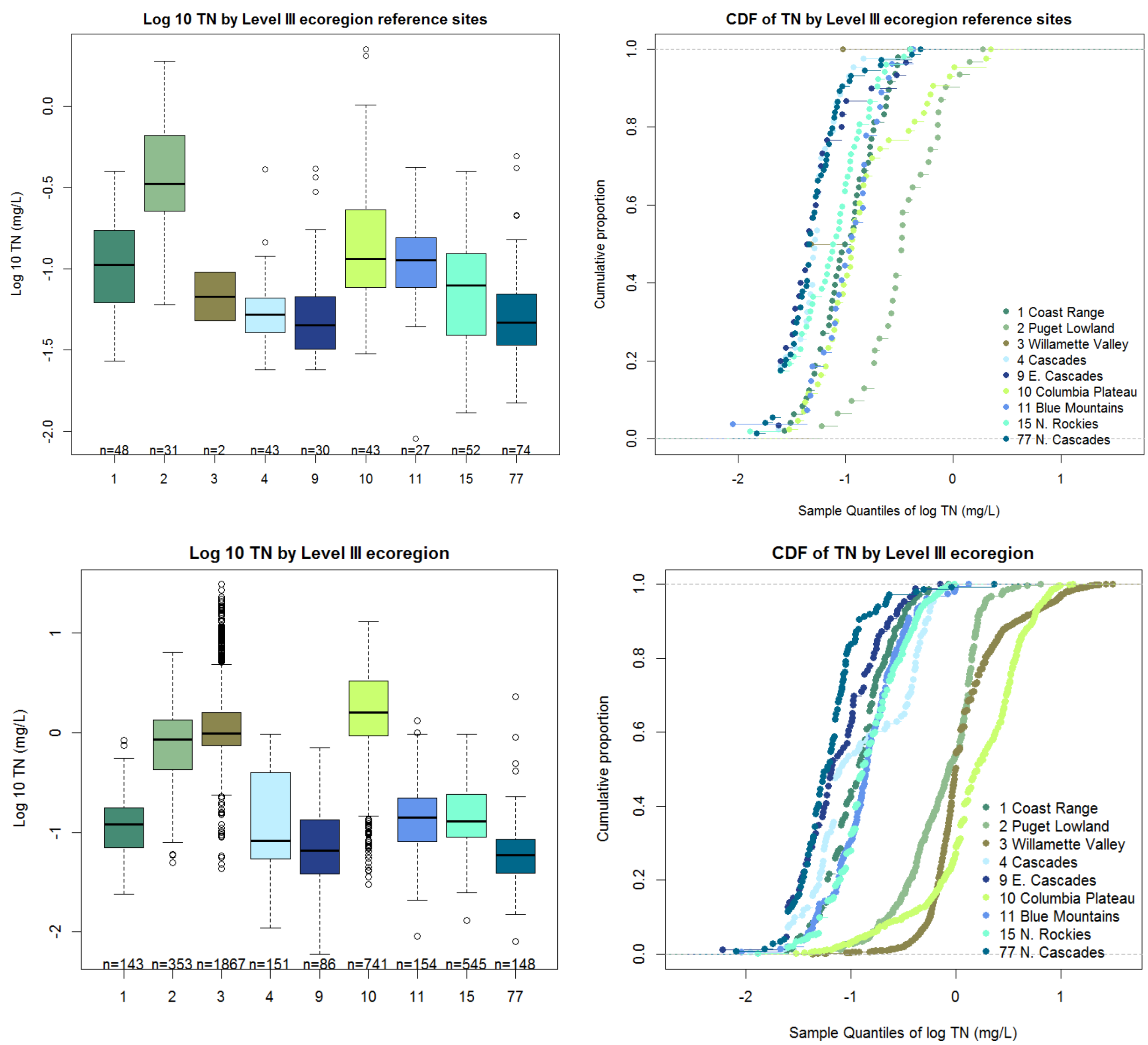


Figure 2 - Boxplots (left) and cumulative distribution functions (right) of log10 transformed TN from reference sampling stations (top) and all sites (bottom) divided by Level III ecoregion. Horizontal lines in cumulative distribution functions indicate tied observation values.

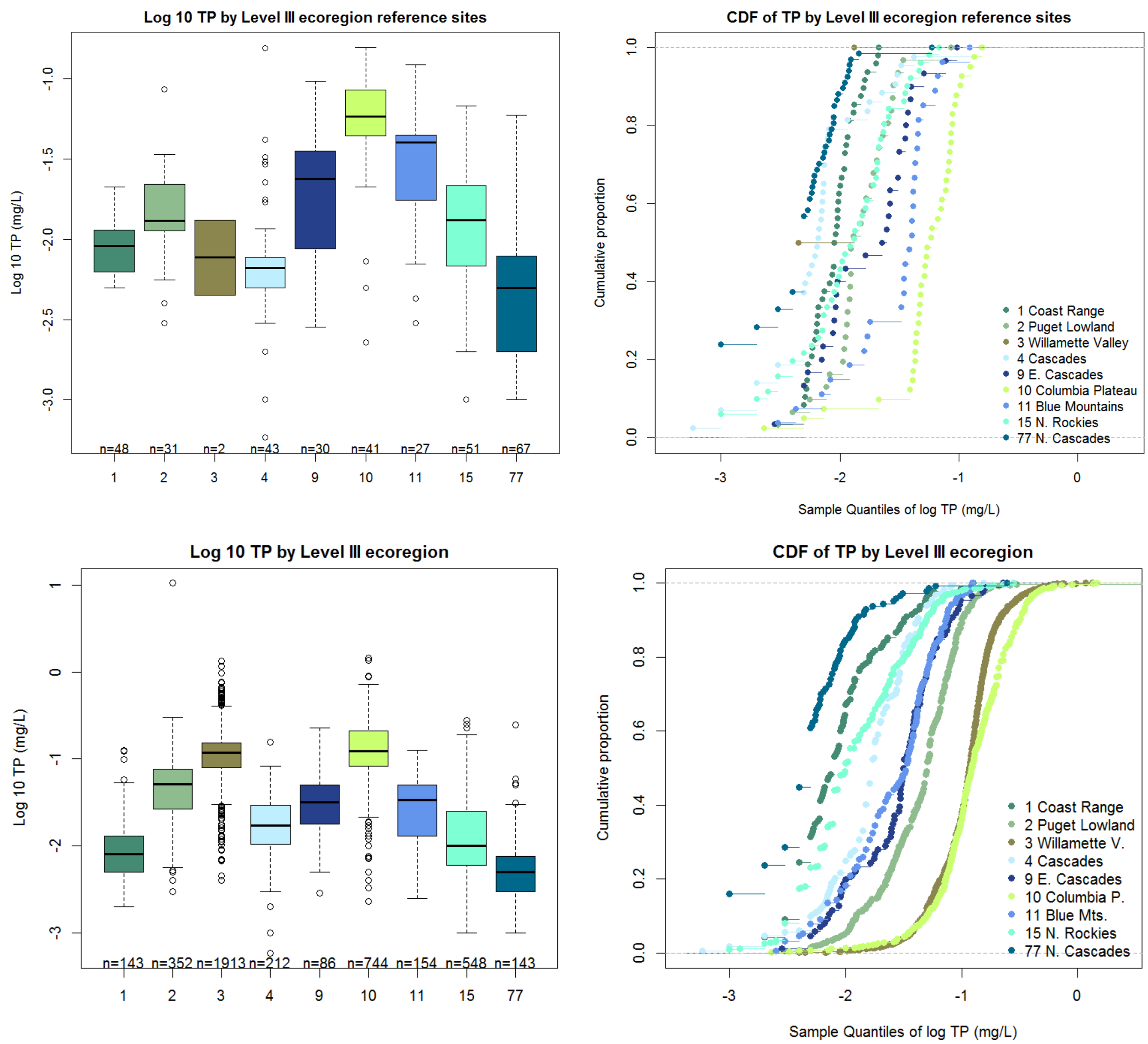


Figure 3 - Boxplots (left) and cumulative distribution functions (right) of log$_{10}$ transformed TP from reference sampling stations (top) and all sites (bottom) divided by Level III ecoregion.

Table 2 - Descriptive statistics of grab sample nutrient and benthic chlorophyll a (Chl a) data for all sites, reference sites and reference sites from ecoregion 2 (Puget Lowlands) from all months. Parameter abbreviations described in Appendix 2.

Analyte	Group	N	Mean	10th	25th	50th	75th	90th
Cl (mg/L)	All	859	1.085	0.241	0.44	0.918	2.989	6.306
	Reference	313	0.693	0.177	0.308	0.53	1.29	4.974
	Reference, ecoregion2	31	2.836	1.1	1.47	1.94	6.405	13.9
TN (mg/L)	All	4188	0.605	0.076	0.221	0.79	1.4	3.2
	Reference	319	0.075	0.025	0.044	0.068	0.12	0.228
	Reference, ecoregion2	31	0.358	0.149	0.226	0.333	0.663	0.854
TP (mg/L)	All	4295	0.054	0.01	0.03	0.08	0.13	0.2
	Reference	309	0.011	0	0.01	0.01	0.03	0.05
	Reference, ecoregion2	31	0.015	0.01	0.01	0.01	0.02	0.03
TSS (mg/L)	All	841	1.952	0.5	1	2	4	8
	Reference	305	1.576	0.5	1	1	3	6
	Reference, ecoregion2	31	2.411	1	1	2	5.5	8
Conductivity (µS/cm)	All	931	88.54	34	54.78	83.9	153	256
	Reference	310	67.599	23.25	39.81	69.34	113.27	182.22
	Reference, ecoregion2	25	98.66	57.6	79.6	89.5	121.15	186.9
DO (mg/L)	All	434	9.616	8.35	9.14	9.8	10.37	10.85
	Reference	217	9.899	8.94	9.4	9.92	10.5	11.06
	Reference, ecoregion2	26	9.9427	9.46	9.8	9.88	10.38	10.61
pH	All	891	7.679	7.1	7.4	7.67	7.99	8.28
	Reference	312	7.608	7.04	7.39	7.63	7.91	8.14
	Reference, ecoregion2	29	7.464	7.1	7.35	7.54	7.69	7.9
Turbidity (NTU)	All	731	0.724	0.1	0.3	0.7	1.8	3.3
	Reference	259	0.57	0.1	0.2	0.5	1.4	2.6
	Reference, ecoregion2	24	0.933	0.3	0.6	1.2	1.9	2.4
Chl a (mg/m^2)	All	626	15.325	3	6	16	41.1	79.8
	Reference	221	12.915	3.8	6.2	12.7	25.2	57.7
	Reference, ecoregion2	28	28.180	9.2	16.7	27.3	46.1	113.2

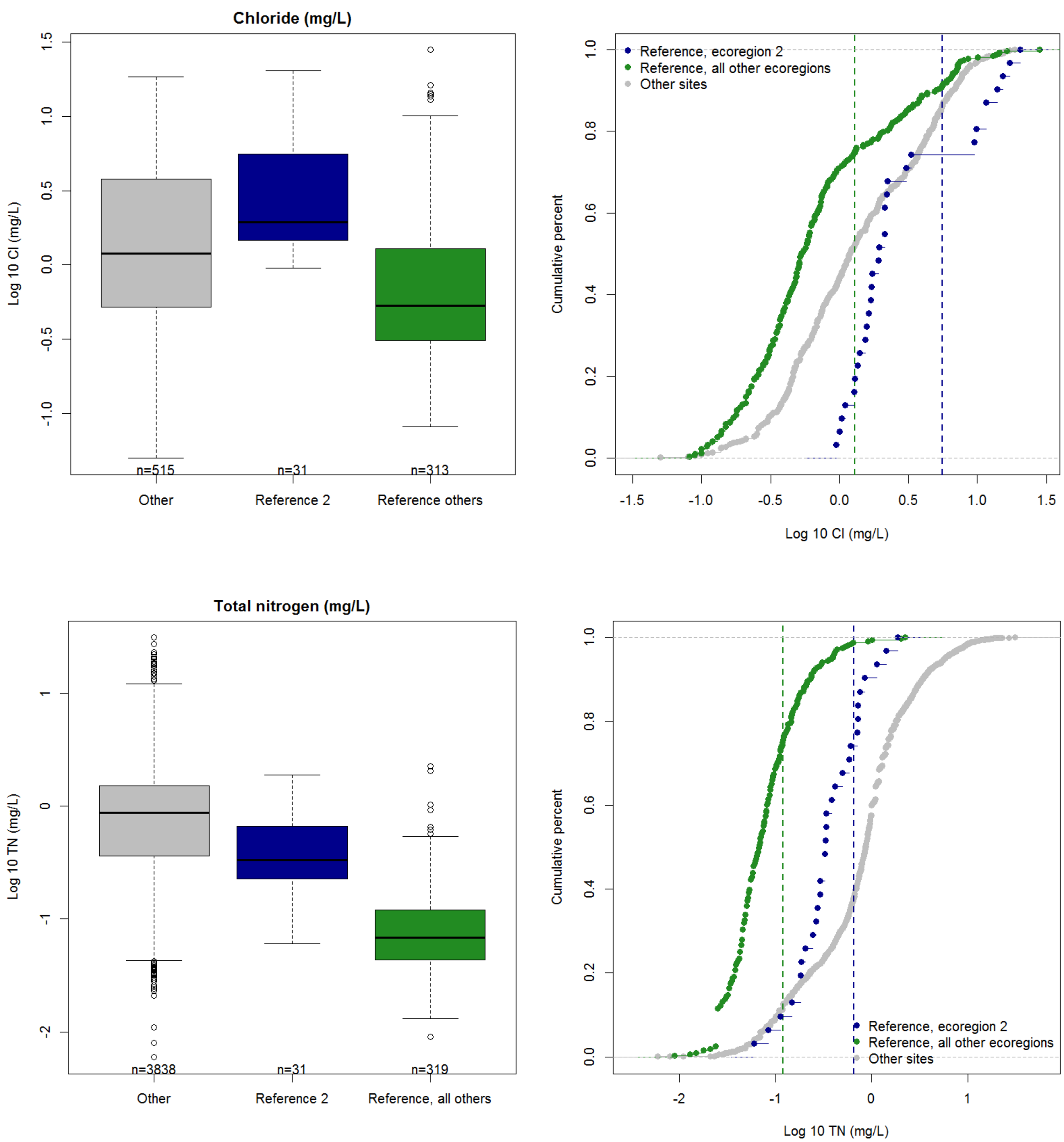


Figure 4 - Distributions of grab sample chloride (top) and TN (bottom) in streams not specifically designated as reference ("Other", grey symbols) and in reference streams, presented separately for Omernik Level III ecoregion 2 (Puget Lowlands, blue symbols) and all other ecoregions (green symbols). The green dashed vertical line is the reference site 75th percentile for non-ecoregion 2 ecoregions. The dark blue dashed vertical lines is the 75th percentile of ecoregion 2. Sample size for these figures are provided in the boxplots and also in Table 2.

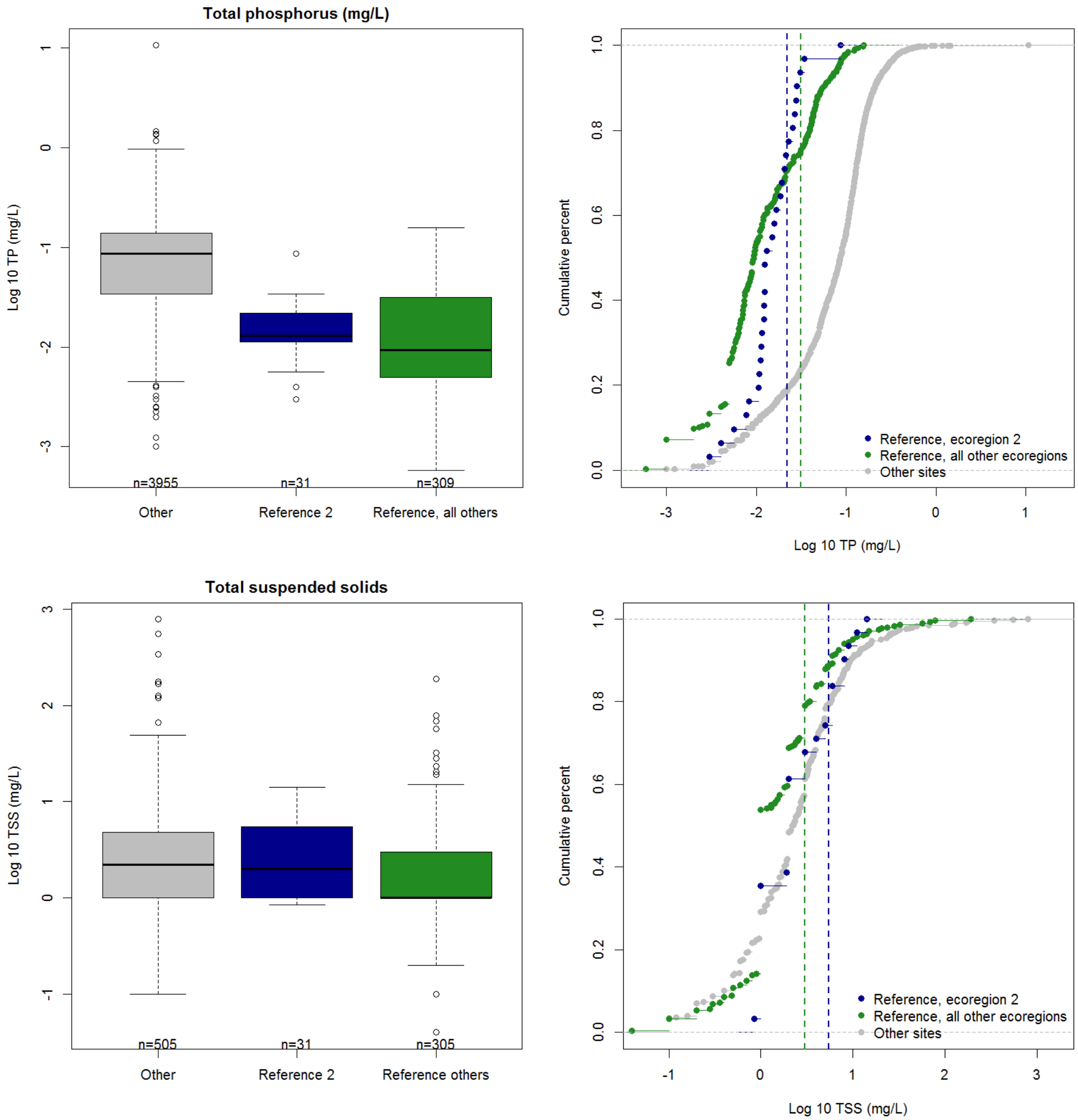


Figure 5 - Distributions of grab sample TP (top) and TSS (bottom) in streams not specifically designated as reference ("Other", grey symbols) and in reference streams, presented separately for Omernik Level III ecoregion 2 (Puget Lowlands, blue symbols) and all other ecoregions (green symbols). The green dashed vertical line is the reference site 75th percentile for non-ecoregion 2 ecoregions. The dark blue dashed vertical lines is the 75th percentile of ecoregion 2. Sample size for these figures are provided in the boxplots and also in Table 2.

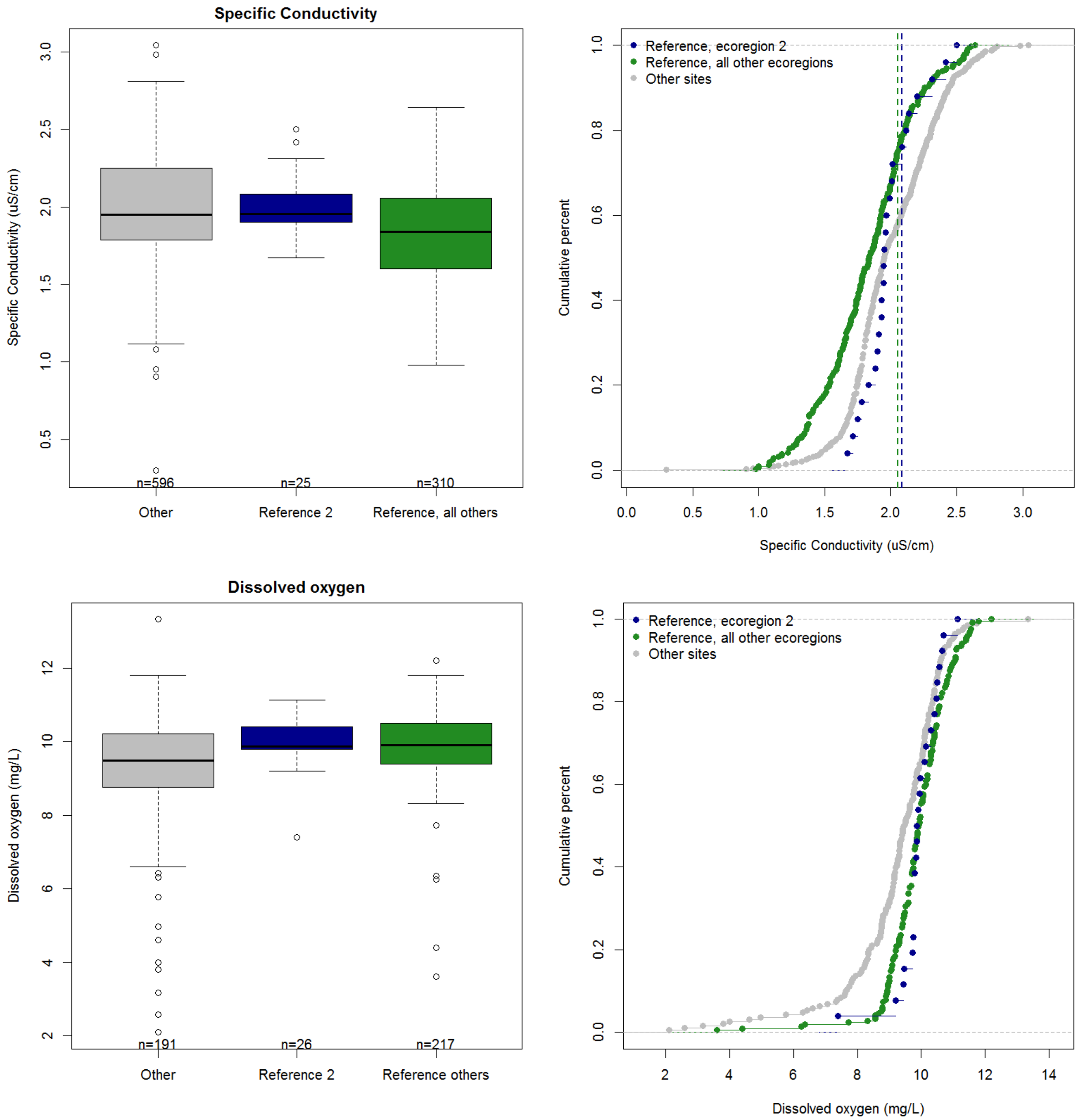


Figure 6 - Distributions of grab sample specific conductivity (top) and dissolved oxygen (bottom) in streams not specifically designated as reference ("Other", grey symbols) and in reference streams, presented separately for Omernik Level III ecoregion 2 (Puget Lowlands, blue symbols) and all other ecoregions (green symbols). The green dashed vertical line is the reference site 75th percentile for non-ecoregion 2 ecoregions. The dark blue dashed vertical lines is the 75th percentile of ecoregion 2. Sample size for these figures are provided in the boxplots and also in Table 2.

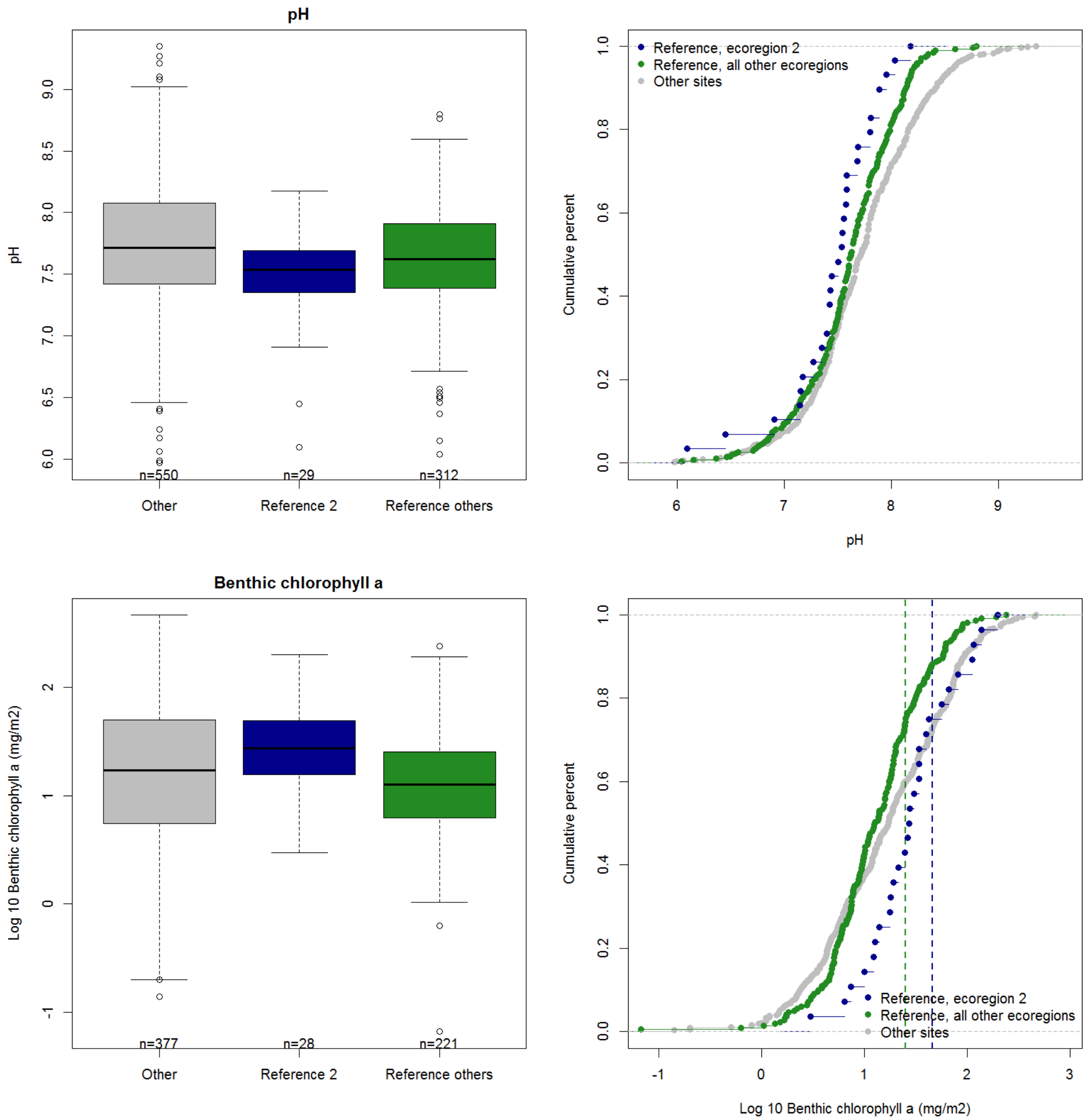


Figure 7 - Distributions of grab sample pH (top) and benthic chlorophyll a (bottom) in streams not specifically designated as reference ("Other", grey symbols) and in reference streams, presented separately for Omernik Level III ecoregion 2 (Puget Lowlands, blue symbols) and all other ecoregions (green symbols). The green dashed vertical line is the reference site 75th percentile for non-ecoregion 2 ecoregions. The dark blue dashed vertical lines is the 75th percentile of ecoregion 2. Sample size for these figures are provided in the boxplots and also in Table 2.

Stressor response

Benthic chlorophyll *a* had no apparent relationship with nutrient concentration (Figure 8). Diatom metrics developed based on diatom nutrient optima, however, responded as expected to nutrient concentrations, increasing with TN and TP respectively (Figure 9 and Figure 10, and see Appendix 4). For the periphyton metrics, reference site quartiles and deciles were calculated as response targets and TN and TP values associated with these targets were interpolated from the simple linear regression models using the TN and TP concentrations at which response target values intersected the mean regression line (Table 3 and Table 4).

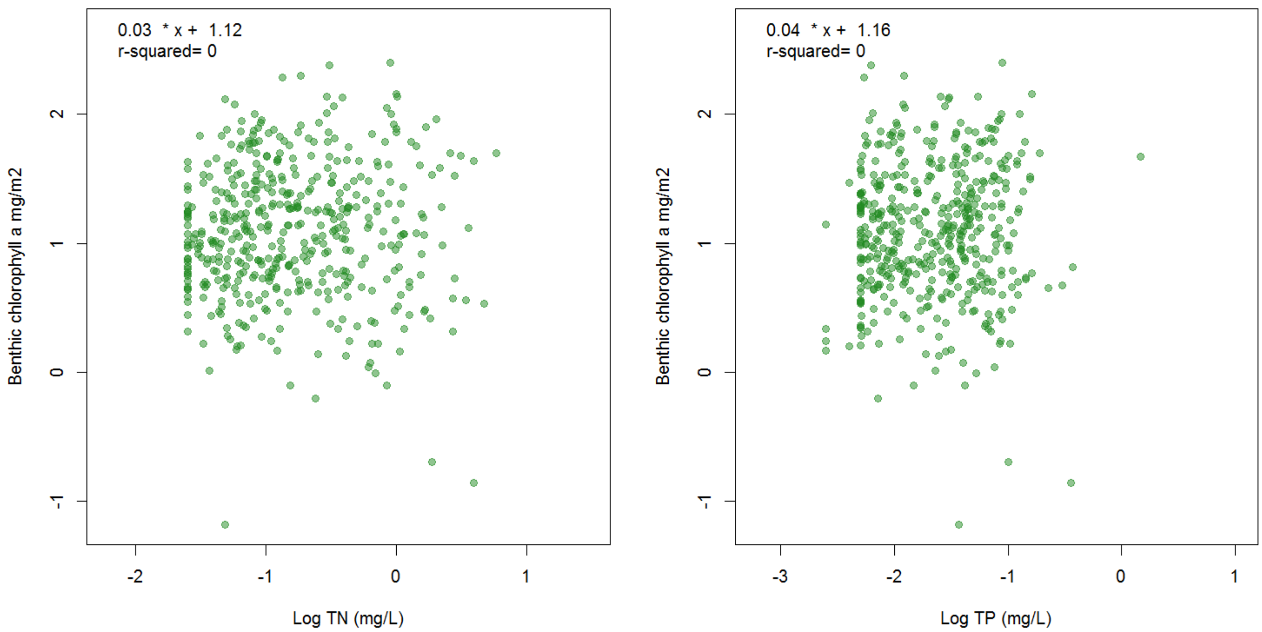


Figure 8 - Benthic chlorophyll *a* (mg/m^2) in relation to grab sample nutrient concentration.

Table 3 - TN endpoints interpolated from periphyton metric regression models as responses to TN for all reference sites. All linear regressions presented were statistically significant ($p < 0.001$) in the proper response direction (ecologically sound). Rho = spearman's correlation coefficient, b = regression intercept, m = regression slope, r^2 = variance explained by regression, q = percentile of deciles (90th) and quartiles (75th), TN90 and TN75 are the interpolated log-base TN values (first value) and back-transformed values (second value) associated with the reference percentile of each metric. Metrics and metric sources are explained in the Methods and in Appendix 3.

Metric	Abbreviation	rho	intercept	slope	r2	q90	TN90		q75	TN75	
TN Optima Index	wa_OptCat_LNtl	0.46	2.73	0.45	0.21	2.68	-0.11	**0.78**	2.28	-1.01	**0.10**
% Disturbed Land Index	wa_OptCat_L1DisTot	0.45	2.90	0.45	0.19	2.81	-0.21	**0.62**	2.49	-0.91	**0.12**
Multivariate Disturbed Land Index	wa_OptCat_DisTotMMI	0.38	2.52	0.48	0.14	2.57	0.10	**1.26**	2.15	-0.77	**0.17**
TN and TP Index	wa_OptCat_NutMMI	0.37	2.49	0.46	0.14	2.56	0.16	**1.45**	2.16	-0.71	**0.19**
Embeddedness Index	wa_OptCat_XEMBED	0.36	2.24	0.42	0.15	2.15	-0.22	**0.60**	1.84	-0.94	**0.11**

Table 4 – TP endpoints interpolated from periphyton metric regression models as responses to TP. All other details as in Table 3.

Metric	Abbreviation	rho	intercept	slope	r2	q90	TP90		q75	TP75	
TP Optima Index	wa_OptCat_L1Ptl	0.67	3.31	0.72	0.41	2.52	-1.10	**0.080**	2.07	-1.72	**0.019**
TN and TP Index	wa_OptCat_NutMMI	0.65	3.36	0.71	0.38	2.57	-1.11	**0.078**	2.18	-1.66	**0.022**
Multivariate Disturbed Land Index	wa_OptCat_DisTotMMI	0.63	3.34	0.69	0.34	2.61	-1.05	**0.089**	2.16	-1.71	**0.020**
Embeddedness Index	wa_OptCat_XEMBED	0.58	2.86	0.56	0.30	2.15	-1.28	**0.052**	1.85	-1.81	**0.015**
Conductivity Index	wa_OptCat_LCond	0.57	3.18	0.61	0.27	2.58	-0.98	**0.106**	2.11	-1.74	**0.018**

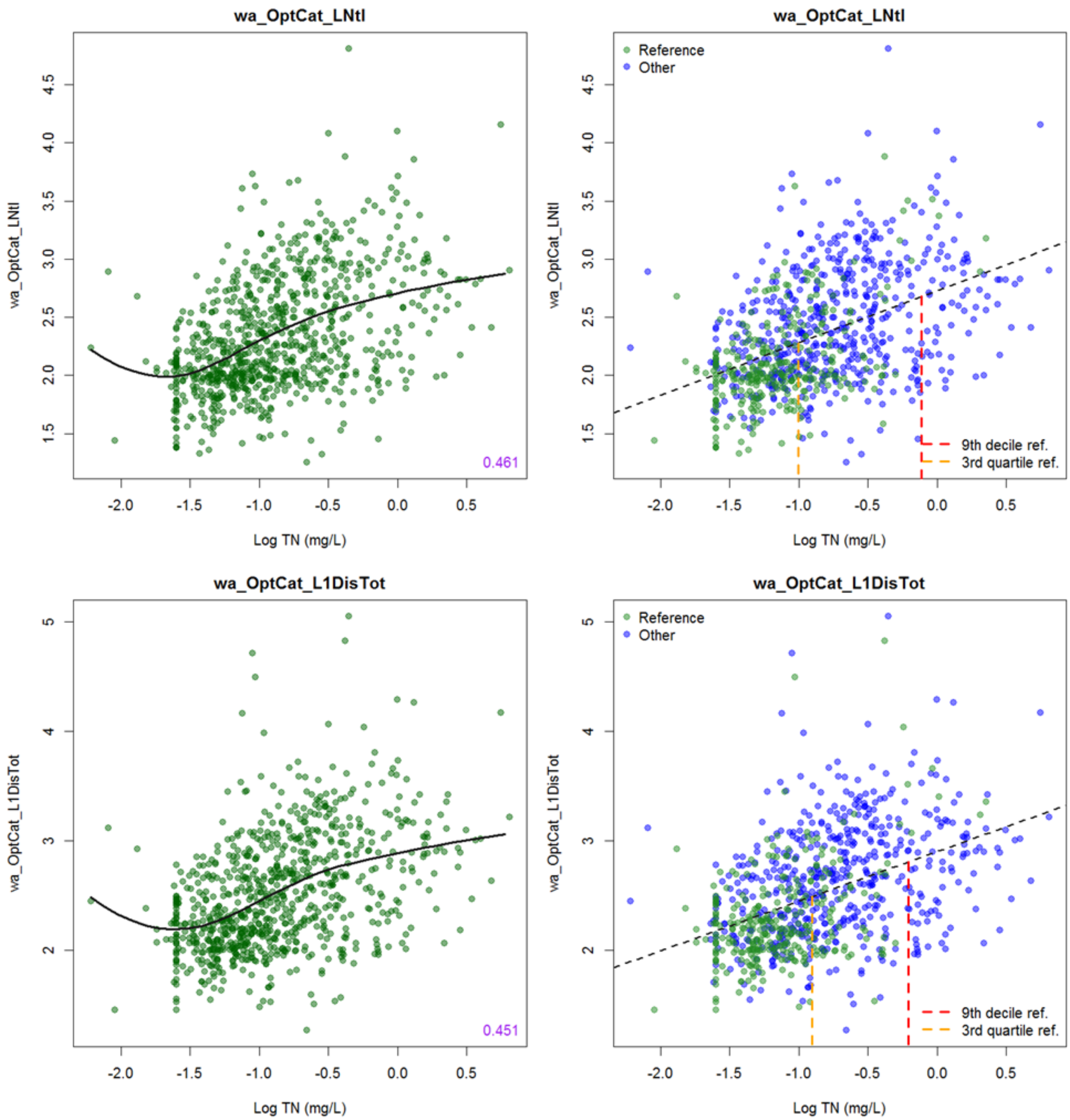


Figure 9 - Two selected diatom metrics shown in relation to TN. Loess lines plots (left) annotated with Spearman's rho (loess span=0.75). Simple linear regression model (right) with the 9th decile (90%) and 3rd quartile (75%) of the reference site metric distribution identified and the nutrient concentrations associated with those response targets indicated with vertical dashed lines (right). For a description of the periphyton metrics, see Methods and Appendix 3.

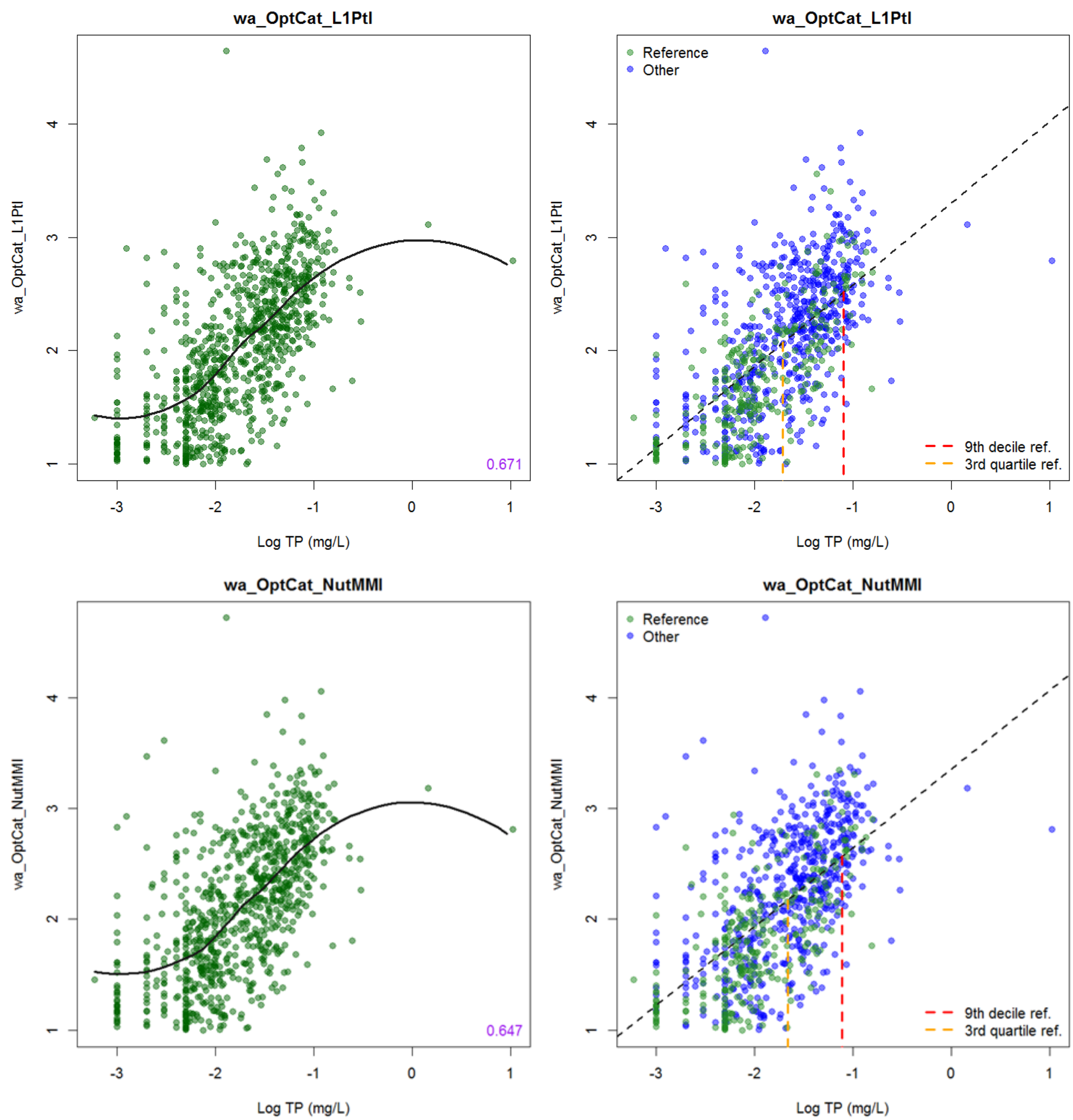


Figure 10 - Two selected diatom metrics shown in relation to TP. Loess lines plots (left) annotated with Spearman's rho (loess span=0.75). Simple linear regression model (right) with the 9th decile and 3rd quartile of the reference site metric distribution identified and the nutrient concentrations associated with those response targets indicated with vertical dashed lines (right). For a description of the periphyton metrics, see Methods and Appendix 3.

Classification:

Residual analysis

For classification, we looked at three independent, supporting lines of analysis: residual analysis, recursive partitioning, and Nash-Sutcliffe efficiency. Residual analysis revealed that diatom metric - nutrient concentration relationships varied along various physical gradients and, as a whole, ecoregions 3 (Willamette Valley), 10 (Columbia Plateau), and 11 (Blue Mountains) tended to have higher biased residuals than other ecoregions for both nutrients, but even more so for TN (Figure 11 and Figure 12). Additional model output in Appendix 5.

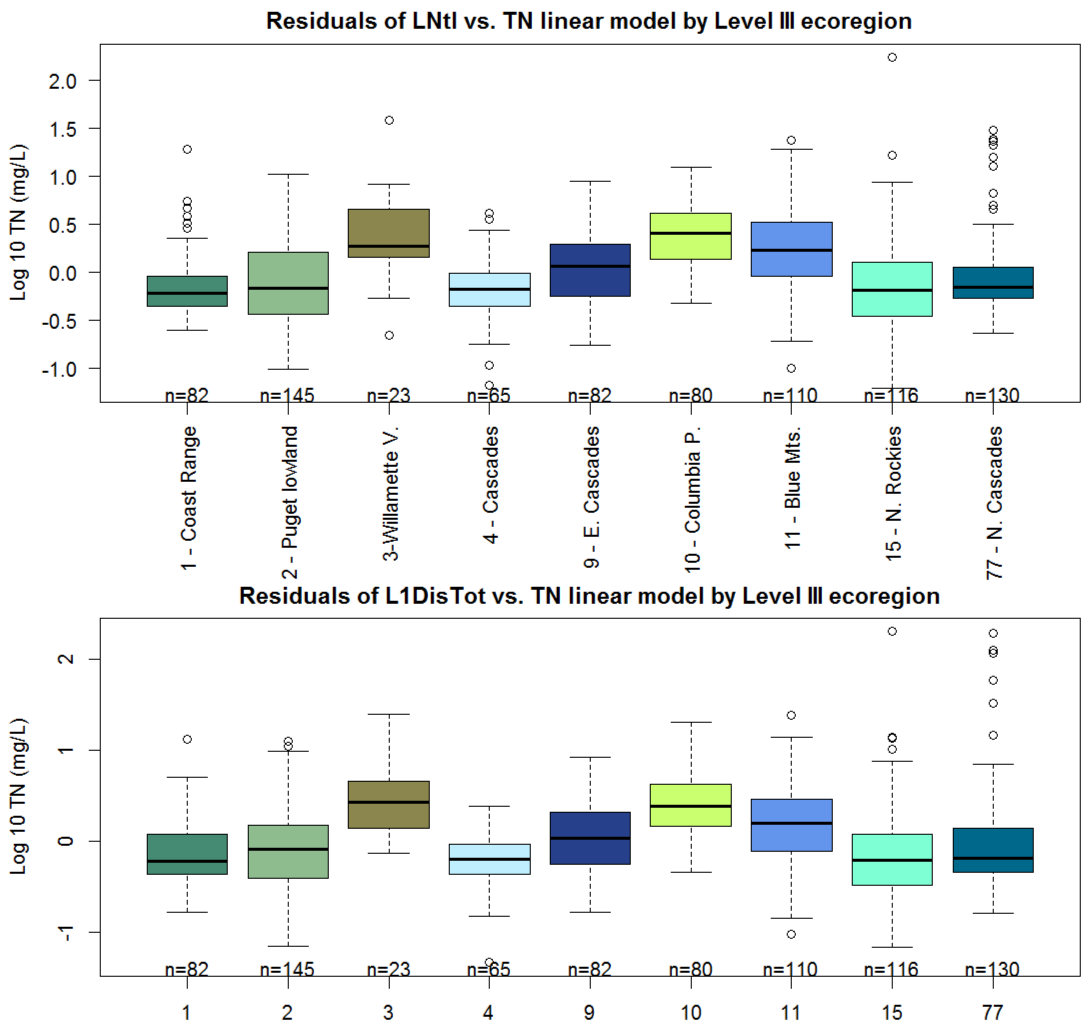


Figure 11 - Residuals of the weighted average LNtl diatom metric (wa_OptCat_LNtl) vs. log_{10} transformed TN linear model (top) and weighted average L1DisTot diatom metric (wa_OptCat_L1DisTot) vs. log_{10} transformed TN model (bottom) as a function of Omernik Level III ecoregion.

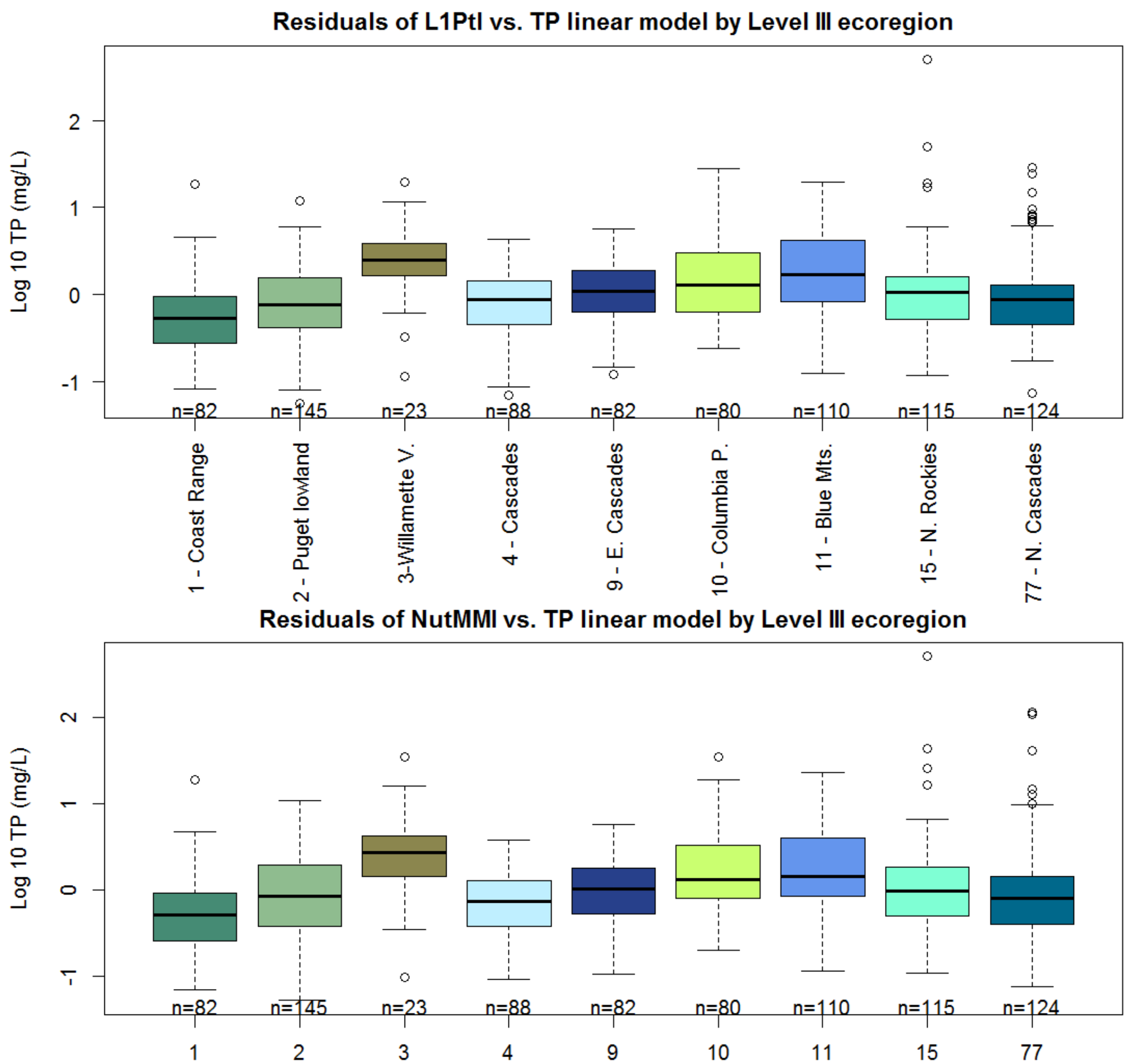


Figure 12 - Residuals of the weighted average L1Ptl diatom metric (wa_OptCat_L1Ptl) vs. $\log_{10}$ transformed TP linear model (top) and weighted average NutMMI diatom metric (wa_OptCat_NutMMI) vs. $\log_{10}$ transformed TP model (bottom) as a function of Omernik Level III ecoregion.

Model-based recursive partitioning

Model based-recursive partitioning for metrics responsive to TN and TP reinforced the differences observed in the residual analysis. When ecoregion was used as the splitting variable, models for both nutrients split with ecoregions 3, 10 and 11 (Willamette Valley, Columbia Plateau, and Blue Mountains, respectively) apart from the rest (Figure 13 and Figure 14). Ecoregion 3, 10, and 11 sites had a steeped metric response to TN (Figure 13). Functional differences were not as great in response to TP (Figure 14). Additional model output shown in Appendix 5.

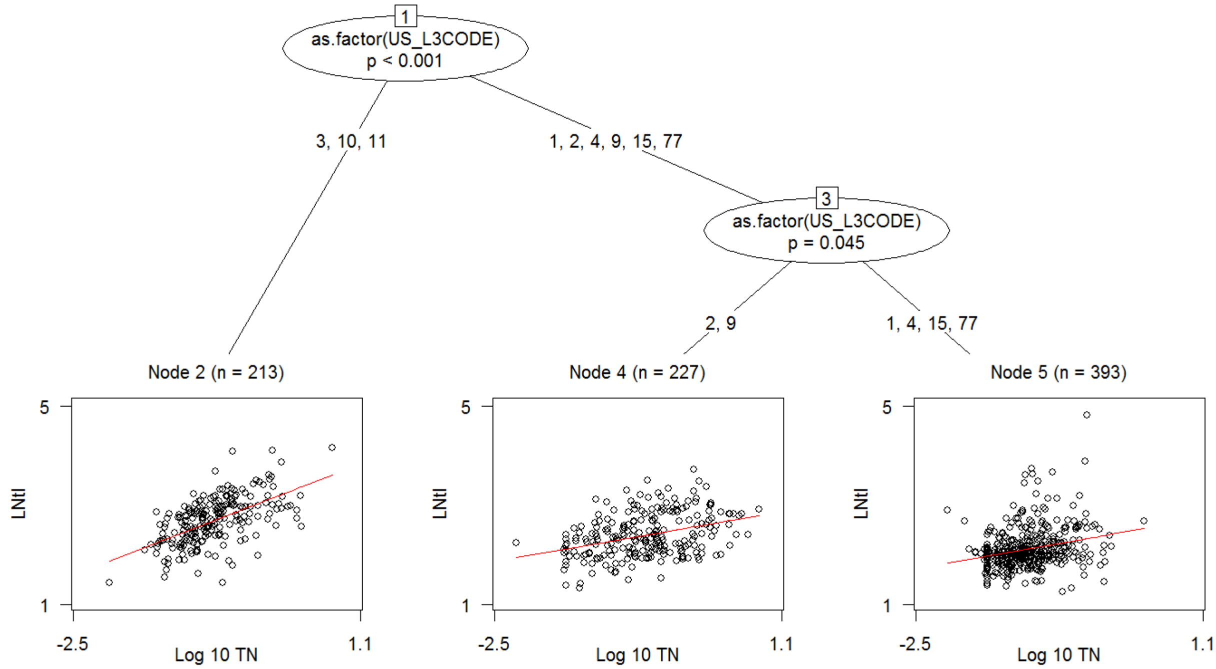


Figure 13 - Model-based recursive partitioning of the wa_OptCat_LNtl optima metric as a response to TN concentration using Omernik Level III ecoregion as the splitting variable.

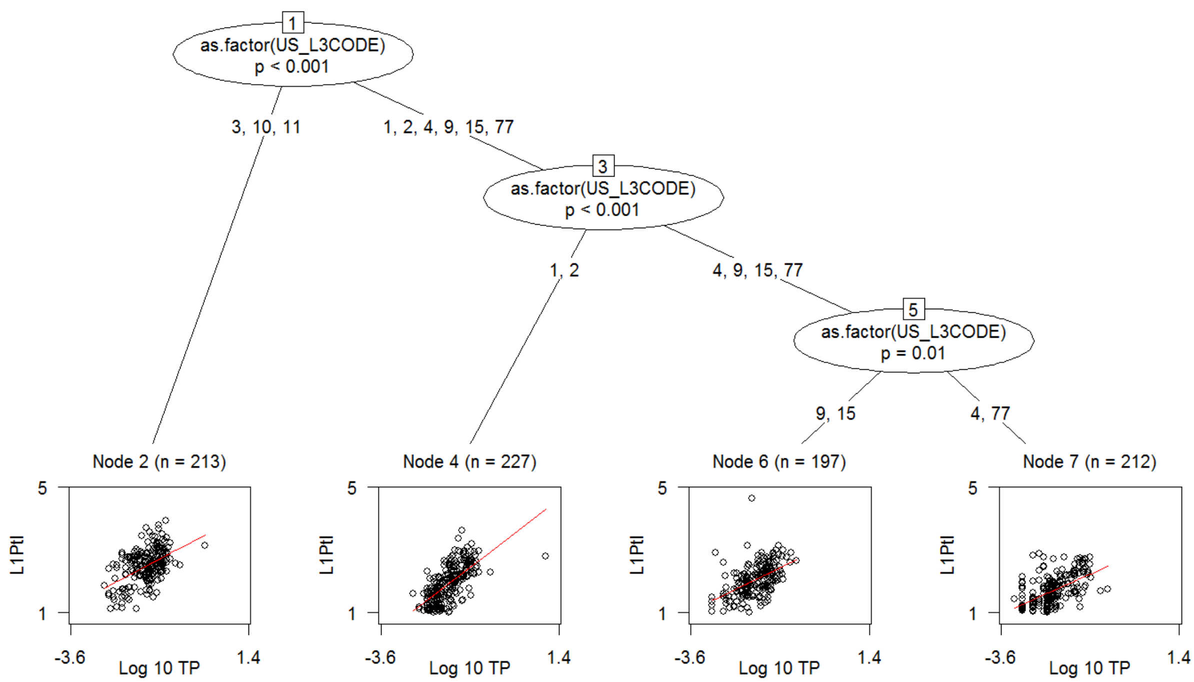


Figure 14 - Model-based recursive partitioning of the wa_OptCat_L1Ptl optima metric as a response to TP concentration using Omernik Level III ecoregion as the splitting variable.

Nash-Sutcliffe efficiency calculation

Finally, the NSE for models built from metrics sensitive to TN and TP concentration using ecoregions 3, 10, and 11 (Willamette Valley, Columbia Plateau, and Blue Mountains, respectively) changed substantially when run for the other ecoregions (Figure 15), supporting a different relationship for these three ecoregions consistent with the other two classification analyses. More detail provided in Appendix 5.

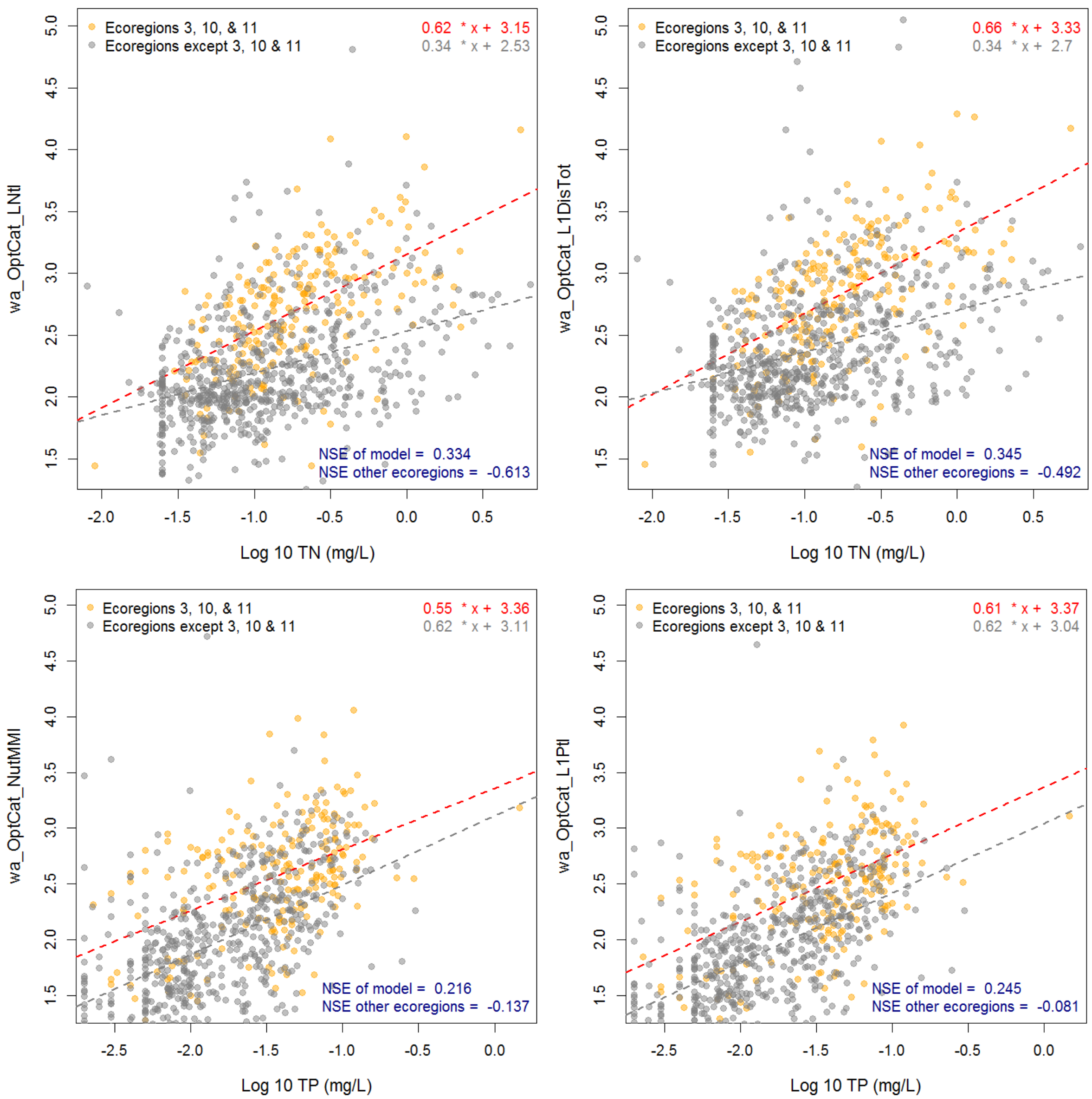


Figure 15 - Diatom metrics from sampling stations from ecoregions 3, 10, and 11 (Willamette Valley, Columbia Plateau, and Blue Mountains, respectively) were used to build models against which models built from samples in all remaining ecoregions were tested. Top performing TN metrics (top) and TP metrics (bottom).

Stressor- response revisited – Ecoregional models

Based on the classification analysis lines of evidence, it was decided to separate ecoregions 3, 10 and 11 (Willamette Valley, Columbia Plateau, and Blue Mountains, respectively) from the remaining Level III ecoregions in Washington State as two separate classes, and stressor-response models were rebuilt for these distinct groups. For TN, linear regression models for ecoregions 3, 10 and 11 have steeper slopes and higher intercepts than the remaining ecoregions (Figure 16 and Figure 17, Table 5 and Table 6). The reasons for this may include greater relative N limitation in ecoregions 3, 10, and 11, but also natural variability in metric response. Again, the fact that the trend is observed across the same range in TN would suggest, potentially, greater P availability in ecoregions 3, 10, and 11 leading to greater sensitivity to N. In support, one can see in the bottom of Figure 15 that ecoregion 3, 10, and 11 sites tend to cluster towards the higher end of TP concentrations. For TP, the metrics in ecoregions 3, 10, and 11 had similar slopes but higher intercepts than the other ecoregions (Figure 18 and Figure 19, Table 7 and Table 8). One could separate sites by ecoregions for TP, but the similarity in TP models across the two classes and their generally lower correlation coefficients as separate classes versus a single class (Table 7 and Table 8 versus Table 4) might suggest using the classification only for TN and not TP.

Table 5 - TN endpoints interpolated from periphyton metric regression models as responses to TN in ecoregions 3, 10, and 11 (Willamette Valley, Columbia Plateau, and Blue Mountains, respectively). All other details as in Table 3.

Metric	Abbreviation	rho	intercept	slope	r2	q90	TN90		q75	TN75	
% Disturbed Land Index	wa_OptCat_L1DisTot	0.60	3.33	0.66	0.35	3.24	-0.13	**0.74**	2.95	-0.59	**0.26**
TN Optima Index	wa_OptCat_LNtl	0.58	3.15	0.62	0.33	2.99	-0.26	**0.55**	2.79	-0.58	**0.27**
Multivariate Disturbed Land Index	wa_OptCat_DisTotMMI	0.53	3.04	0.62	0.27	2.97	-0.13	**0.75**	2.76	-0.45	**0.35**
TN and TP Index	wa_OptCat_NutMMI	0.51	2.98	0.56	0.23	2.95	-0.05	**0.90**	2.74	-0.41	**0.39**
Embeddedness Index	wa_OptCat_XEMBED	0.39	2.54	0.48	0.18	2.58	0.08	**1.19**	2.36	-0.39	**0.41**

Table 6 - TN endpoints interpolated from periphyton metric regression models as responses to TN in ecoregions 1, 2, 4, 9, 15, and 77 (Coast Range, Puget Lowland, Cascades, Eastern Cascades, Northern Rockies, and North Cascades, respectively). All other details as in Table 3.

Metric	Abbreviation	rho	intercept	slope	r2	q90	TN90		q75	TN75	
% Disturbed Land Index	wa_OptCat_LNtl	0.40	2.53	0.34	0.15	2.34	-0.57	**0.27**	2.13	-1.19	**0.06**
TN Optima Index	wa_OptCat_L1DisTot	0.38	2.70	0.34	0.13	2.56	-0.41	**0.39**	2.32	-1.13	**0.07**
Multivariate Disturbed Land Index	wa_OptCat_DisTotMMI	0.29	2.26	0.35	0.09	2.20	-0.18	**0.66**	1.96	-0.86	**0.14**
TN and TP Index	wa_OptCat_NutMMI	0.28	2.24	0.34	0.09	2.22	-0.08	**0.83**	1.98	-0.78	**0.17**
Embeddedness Index	wa_OptCat_XEMBED	0.28	2.08	0.35	0.11	1.97	-0.31	**0.49**	1.71	-1.06	**0.09**

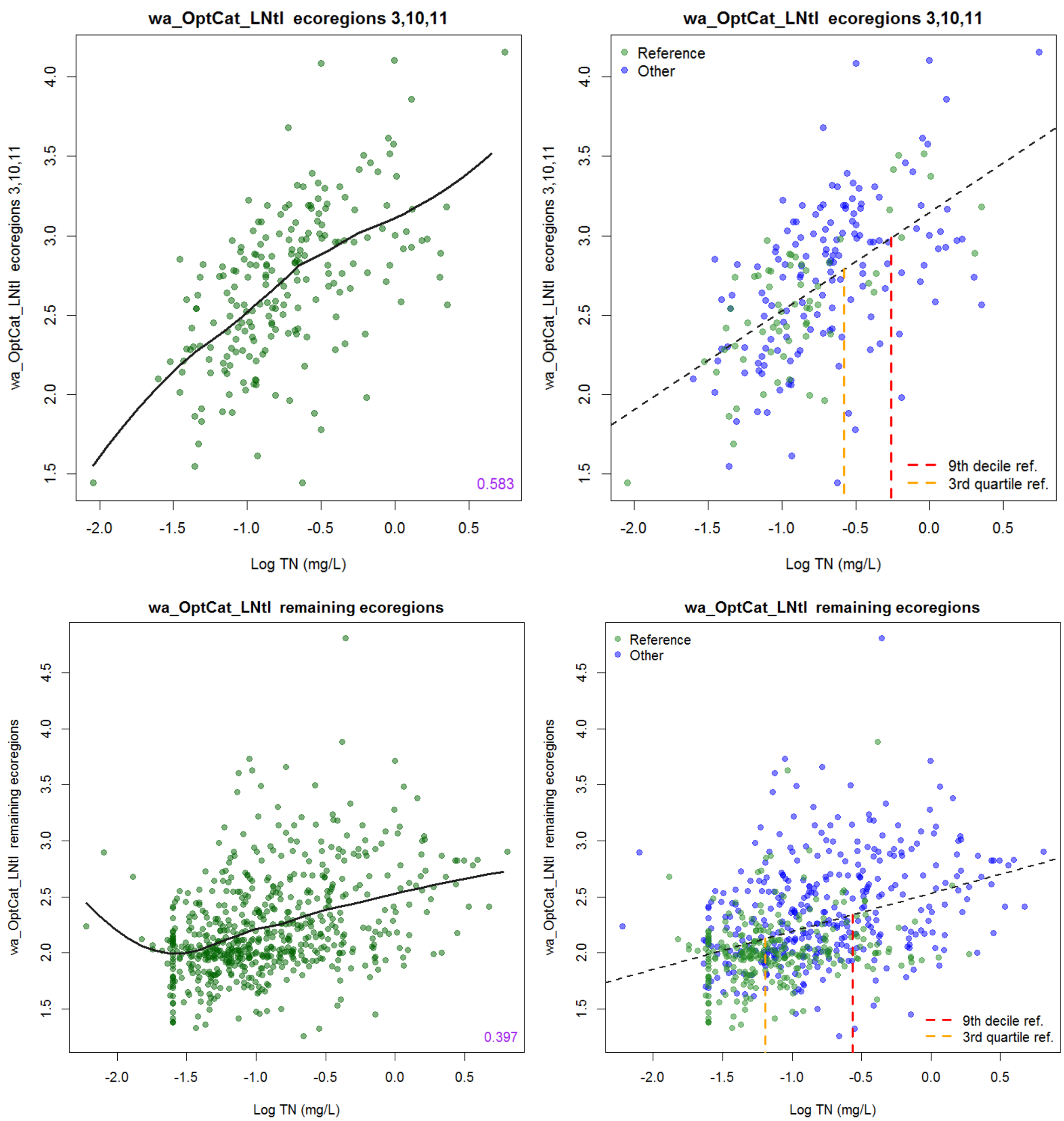


Figure 16 - Diatom metric wa_OptCat_L1Ntl in relation to TN in samples from ecoregions 3, 10 and 11 ((Willamette Valley, Columbia Plateau, and Blue Mountains, respectively, top) and the remaining ecoregions (bottom). Loess lines plots (left) annotated with Spearman's rho (loess span=0.75). Simple linear regression model (right) with the 9th decile and 3rd quartile of the reference site metric distribution identified and the nutrient concentrations associated with those response targets represented as dashed lines (right). For a description of the periphyton metrics, see Methods and Appendix 3.

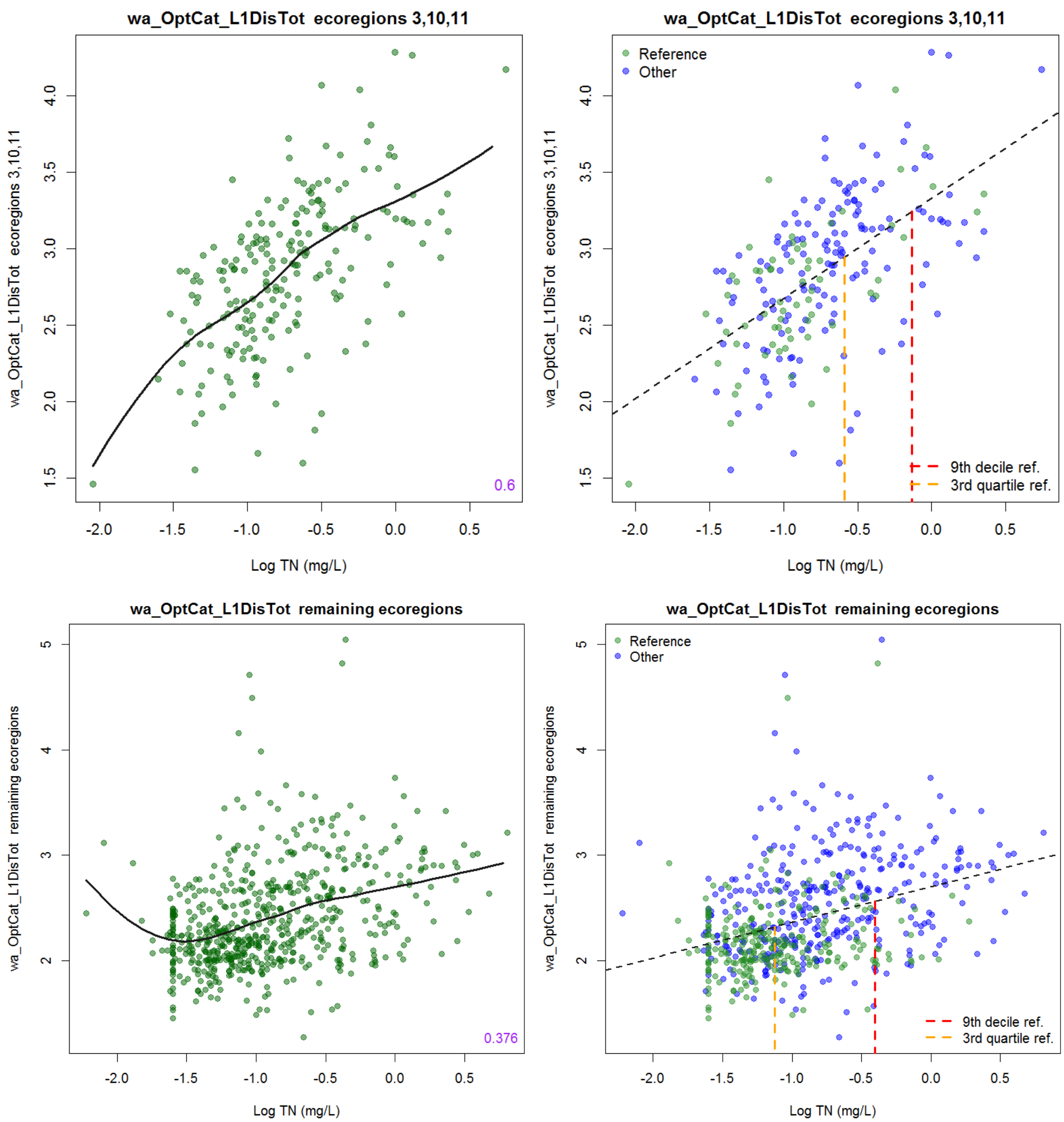


Figure 17 - Diatom metric wa_OptCat_DisTot in relation to TN in samples from ecoregions 3, 10 and 11 (Willamette Valley, Columbia Plateau, and Blue Mountains, respectively, top) and the remaining ecoregions (bottom). Loess lines plots (left) annotated with Spearman's rho (loess span=0.75). Simple linear regression model (right) with the 9th decile and 3rd quartile of the reference site metric distribution identified and the nutrient concentrations associated with those response targets represented as dashed lines (right). For a description of the periphyton metrics, see Methods and Appendix 3.

Table 7 - TP endpoints interpolated from periphyton metric regression models as responses to TP in ecoregions 3, 10, and 11 (Willamette Valley, Columbia Plateau, and Blue Mountains, respectively). All other details as in Table 3.

Metric	Metric name	rho	intercept	slope	r2	q90	TP90		q75	TP75	
TN and TP Index	wa_OptCat_NutMMI	0.45	3.36	0.55	0.22	2.95	-0.75	**0.18**	2.74	-1.12	**0.075**
TP Optima Index	wa_OptCat_L1Ptl	0.44	3.37	0.61	0.25	2.95	-0.69	**0.20**	2.68	-1.14	**0.072**
Multivariate Disturbed Land Index	wa_OptCat_DisTotMMI	0.43	3.36	0.53	0.19	2.97	-0.74	**0.18**	2.76	-1.12	**0.076**
Embeddedness Index	wa_OptCat_XEMBED	0.31	2.68	0.34	0.09	2.58	-0.31	**0.49**	2.36	-0.97	**0.110**
Conductivity Index	wa_OptCat_LCond	0.29	3.10	0.36	0.08	3.03	-0.19	**0.65**	2.72	-1.07	**0.085**

Table 8 - TP endpoints interpolated from periphyton metric regression models as responses to TP in ecoregions 1, 2, 4, 9, 15, and 77 (Coast Range, Puget Lowland, Cascades, Eastern Cascades, Northern Rockies, and North Cascades, respectively). All other details as in Table 3.

Metric	Metric name	rho	intercept	slope	r2	q90	TP90		q75	TP75	
TP Optima Index	wa_OptCat_L1Ptl	0.65	3.07	0.63	0.37	2.18	-1.40	**0.040**	1.92	-1.81	**0.016**
TN and TP Index	wa_OptCat_NutMMI	0.62	3.13	0.63	0.33	2.22	-1.44	**0.036**	1.98	-1.82	**0.015**
Multivariate Disturbed Land Index	wa_OptCat_DisTotMMI	0.60	3.09	0.61	0.30	2.21	-1.47	**0.034**	1.98	-1.85	**0.014**
Embeddedness Index	wa_OptCat_XEMBED	0.57	2.78	0.54	0.29	1.97	-1.50	**0.032**	1.71	-1.97	**0.011**
Conductivity Index	wa_OptCat_LCond	0.55	2.92	0.53	0.24	2.20	-1.35	**0.044**	1.96	-1.82	**0.015**

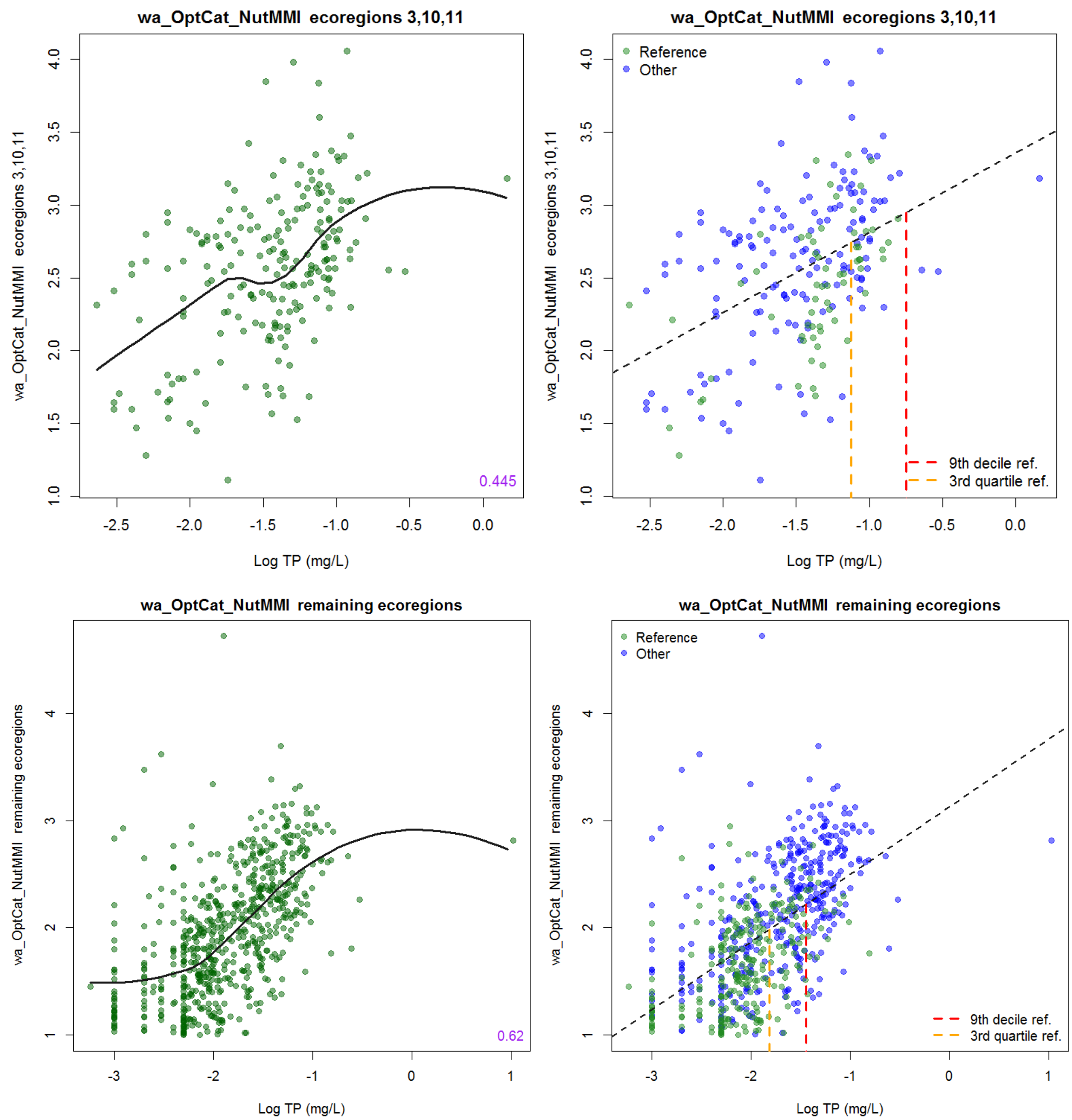


Figure 18 - Diatom metric wa_OptCat_NutMMI in relation to TP in samples from ecoregions 3, 10 and 11 (Willamette Valley, Columbia Plateau, and Blue Mountains, respectively, top) and the remaining ecoregions (bottom). Loess lines plots (left) annotated with Spearman's rho (loess span=0.75). Simple linear regression model (right) with the 9th decile and 3rd quartile of the reference site metric distribution identified and the nutrient concentrations associated with those response targets represented as dashed lines (right). For a description of the periphyton metrics, see Methods and Appendix 3.

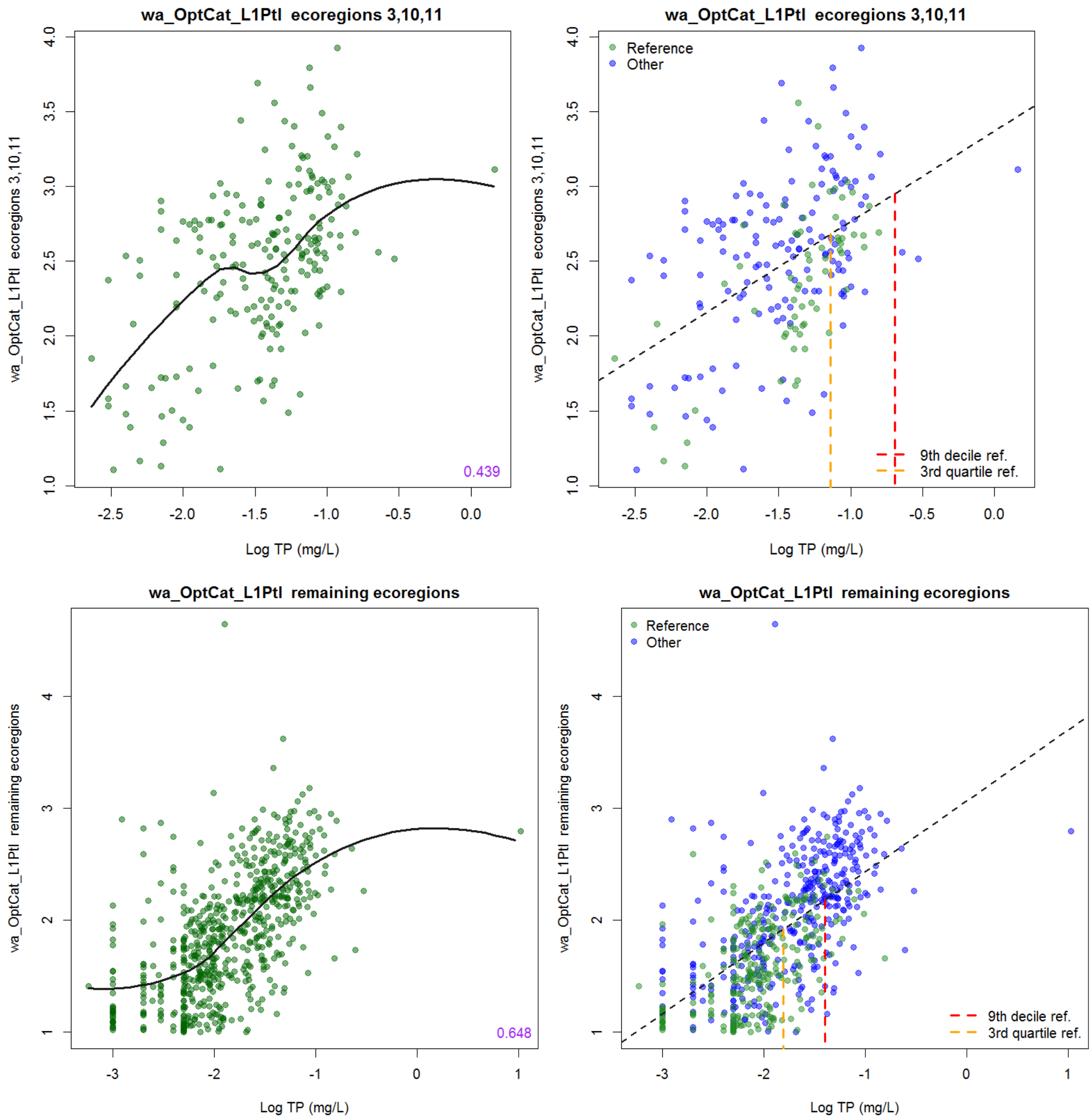


Figure 19 - Diatom metric wa_OptCat_L1Ptl in relation to TP in samples from ecoregions 3, 10 and 11 (Willamette Valley, Columbia Plateau, and Blue Mountains, respectively, top) and the remaining ecoregions (bottom). Loess lines plots (left) annotated with Spearman's rho (loess span=0.75). Simple linear regression model (right) with the 9th decile and 3rd quartile of the reference site metric distribution identified and the nutrient concentrations associated with those response targets represented as dashed lines (right). For a description of the periphyton metrics, see Methods and Appendix 3.

4. DISCUSSION

General Observations

Periphyton metrics responded to stream nutrient concentration gradients in Washington. As suggested before in many contexts, diatom assemblage structure is a sensitive indicator of nutrient gradients. This is likely not surprising but it is important to demonstrate given the anecdotes one hears about algae not being sensitive to nutrients in streams. Algal composition is especially sensitive to nutrients in Washington streams. Algal biomass can be difficult to precisely characterize at the reach scale and often shows highly variable responses to nutrients, which is likely the source of skepticism over the use of algal indicators in streams. Indeed, in Washington, chlorophyll *a* showed little directional response to nutrient gradients (Figure 8). There are many reasons for this, methodological and ecological. However, the diatom assemblage composition data exhibited a consistent and significant response to both total N and P (e.g., Figure 9 and Figure 10).

The significance of this observation is that algal assemblage tools have now demonstrated their potential for many applications in Washington. The responsiveness to nutrients, especially above reference conditions, means they can be used, at a minimum, as screening tools or confirmatory tools of nutrient pollution and potentially other related water quality impairments (e.g., dissolved oxygen, pH, etc.). Either the values of nutrients derived from the stressor-response analyses above, or the reference based periphyton metric values could be applied in this capacity.

This analysis also supports the application of algal assemblage structure as an additional biological indicator for aquatic life use. EPA recommends the use of algae in biological assessment, along with macroinvertebrates and fish, to provide a complete assessment of biological condition (USEPA 2011). A primary requirement of any biological indicator is that it be responsive to disturbance. This study provides evidence that various diatom metrics respond to nutrient disturbance, indicating the utility of using these assemblages for detecting disturbance and resolving a wide range of nutrient pollution conditions.

Another noteworthy observation was the quality of the stressor-response relationships. Most of the periphyton metric – nutrient response relationships examined were significant. Also, for many stressor-response relationship models, nutrient concentrations explained more than a third of the variability in metric scores, which is relatively high in comparison to the chlorophyll *a* – nutrient relationships or AFDM – nutrient relationships, and like that observed in other studies of periphyton metrics (e.g., Stevenson et al. 2008).

Application

A few potential applications are recommended by this analysis. First, the cumulative distribution of endpoint values derived from the stressor-response relationship models provide reasonable TN and TP screening values consistent with the protection of reference quality periphyton conditions in streams in Washington, as do the periphyton metric reference percentiles themselves. These nutrient distributions, though, indicate a range of nutrient endpoints associated with different responses from the most significant, sensitive periphyton indicators, with alterations outside of the reference condition occurring as low as 0.010 mg/L TP and 0.100 mg/L TN, to less sensitive indicators with higher nutrient values. The ranges are within those reported from other studies (Table 9). ECY will need to decide where along these distributions a screening value should be located. The use of the central tendency of endpoints is easy, but would result in a negative impact to half the periphyton diversity measures. In identifying a protective metric, some exploration of the individual metrics themselves is warranted and a combination of statistical and ecological reasoning should be applied (e.g., identifying if there are specific conditions/considerations and algal attributes of greater concern). The state is encouraged to consider the strength of the relationships, level of significance, and issues like the quality of reference sites in each region if and as they consider deriving specific numeric values from these results.

These periphyton metrics are likely not ready to be used as universal standalone biological indicators, even though they come from peer-reviewed products and are well established (i.e., akin to the EPT richness score for macroinvertebrates). ECY would likely want to vet them through the full bioindicator development process they use for macroinvertebrates or fish (e.g., additional data preparation, metric exploration, metric scoring and testing, metric combination, etc.). . However, they show great promise in this regard for Washington. Washington ECY adding another assemblage (algae) for routine monitoring to characterize biological integrity adds greater depth and breadth to assessment of stressors, and other entities throughout the state are encouraged to sample additional assemblages (USEPA 2011). Outside of assessing aquatic life use, however, given their well-established status, these metrics and the nutrient values associated with them here, are likely valuable additions to other site-specific data and explanatory variables in evaluating reasonable potential, TMDL monitoring and modeling, etc.

Table 9. Summary of literature nutrient thresholds associated with various diatom metrics endpoints. From Evans-White et al. 2013.

Dependent variable	Method	TN	TP	Citation
Diatom nutrient index	regression tree	1.169	0.072	Robertson et al., 2006b
Diatom siltation index	regression tree	0.872	0.074	Robertson et al., 2006b
Diatom biotic index	regression tree	1.169	0.072	Robertson et al., 2006b
Number of diatom taxa	nCPA	NA	0.011	Stevenson et al., 2008
Diatom evenness	nCPA	NA	0.019	Stevenson et al., 2008
Proportion of native diatom taxa	nCPA	NA	0.011	Stevenson et al., 2008
Proportion of low-P native taxa	nCPA	NA	0.018	Stevenson et al., 2008
Diatom species similarity to reference	nCPA	NA	0.026	Stevenson et al., 2008
Low-P diatom individuals, %	nCPA	NA	0.018	Stevenson et al., 2008
High-P diatom individuals, %	nCPA	NA	0.011	Stevenson et al., 2008
Fine-grained depositional substrate				
Abundance of pollution tolerant diatoms, %	regression	0.86	0.280	Black et al., 2011
Alkalophilus diatom richness	regression	NS††	0.050	Black et al., 2011
Abundance of pollution-sensitive diatoms, %	regression	NS	0.090	Black et al., 2011
Abundance of high-TN diatoms, %	regression	0.61	0.060	Black et al., 2011
Abundance of high-TP diatoms, %	regression	0.71	0.060	Black et al., 2011
Abundance of N heterotrophs, %	regression	1.5	0.100	Black et al., 2011
Abundance of motile algae, %	regression	0.27	0.060	Black et al., 2011
Richness of motile algae, %	regression	1.49	0.090	Black et al., 2011

Dependent variable	Method	TN	TP	Citation
Coarse-grained substrate (rock or wood)				
Alkalophilus diatom richness	regression	1.25	0.030	Black et al., 2011
Abundance of high TN diatoms, %	regression	1.45	0.070	Black et al., 2011
Abundance of high-TP diatoms, %	regression	1.3	0.080	Black et al., 2011
Abundance of N heterotrophs, %	regression	0.59	0.130	Black et al., 2011
Abundance of motile algae, %	regression	NS	0.200	Black et al., 2011
Richness motile algae, %	regression	1.79	0.070	Black et al., 2011

Recommended Additional Analyses

There are several next steps ECY could take and the following are just a few suggestions. First, revisiting the classification is recommended. Algal assemblages could be structured by some of the common gradients influential in constraining species distributions – like temperature, pH, flow, etc. We just observed little structure with regards to these drivers, but it is recommended that this be pursued somewhat further, possibly using reduced species lists (i.e., removing rare and common taxa).

Testing the relationships observed here with independent data, either collected specifically to test the gradients or by accessing additional periphyton data, is also recommended; however, this analysis represents the third application of this approach regionally, with comparable results. While this analysis encompassed several different programs, a broad nutrient gradient, and is likely robust, testing the responses in subwatersheds, local regions, or along known nutrient gradients within a small region may help strengthen confidence in the applicability.

Another additional effort that could be considered is an attempt to tie periphyton metric values to other valued ecological endpoints, such as nuisance algal levels (e.g., as defined with user perception studies), dissolved oxygen conditions, or invertebrate conditions. Certainly, one would expect periphyton to influence other assemblages, but linking different levels of periphyton metric values to levels of invertebrate responses would potentially help strengthen linkages to higher trophic levels. Similarly, dissolved oxygen concentrations and diel swings are pathways through which excess algal biomass is known to affect aquatic ecosystems. The degree to which linkages between DO and algal metrics could be made would also help strengthen this linkage and help screen sites for potential impairments and verify TMDL modeling of pollutant sources. Some metrics are designed for DO sensitivity, so there is a presumed linkage, but this could be tested in Washington as well.

A last step to be considered is inference modeling. Many diatoms have had nutrient optima calculated for them and/or that effort could be done for taxa in Washington. Once identified, transfer coefficients can be used to infer the true average nutrient concentrations present at a particular site. Indeed, this has been the approach New Jersey has used for their diatom index (Ponader et al. 2008). Quite often, grab samples or even averages from baseflow grab samples, are poor representations of average nutrient conditions. Diatoms may provide a more accurate picture. This may be especially useful in developing or monitoring targets.

5. REFERENCES

Alexander, W. P., and S. D. Grimshaw. 1996. Treed regression. Journal of Computational and Graphical Statistics 5:156–175.

Evans-White, M. A., B. E. Haggard, and J. T. Scott. 2013. A review of stream nutrient criteria development in the United States. Journal of environmental quality 42 (4): 1002-1014.

Hothorn, T. and A. Zeileis. 2014. partykit: A Modular Toolkit for Recursive Partytioning in R. Working Paper 2014-10. Working Papers in Economics and Statistics, Research Platform Empirical and Experimental Economics, Universitaet Innsbruck. URL http://EconPapers.RePEc.org/RePEc:inn:wpaper:2014-10

Nash, J. E., and J. V. Sutcliffe. 1970. River flow forecasting through conceptual models: Part 1. A discussion of principles. *J. Hydrology* 10(3): 282-290.

Oksanen, J., F. G. Blanchet, R. Kindt, P. Legendre, P. R. Minchin, R. B. O'Hara, G. L. Simpson, Peter Solymos, M. H. H. Stevens and H. Wagner. 2015. vegan: Community Ecology Package. R package version 2.2-1. http://CRAN.R-project.org/package=vegan

PRISM Climate Group, Oregon State University, http://prism.oregonstate.edu , raster data created July 10 2012.

Ponader, K.C., D.F. Charles, T.J Belton, and D.M. Winter. 2008. Total phosphorus inference models and indices for coastal plain streams based on benthic diatom assemblages from artificial substrates. Hydrobiologia 610:139–152.

R Core Team (2016). R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. URL https://www.R-project.org/

Stevenson, R.J., Y. Pan, K.M. Manoylov, C.A. Parker, D. P. Larsen and A.T. Herlihy. 2008. Development of diatom indicators of ecological conditions for streams of the western US. Journal of the North American Benthological Society 27: 1000-1016.

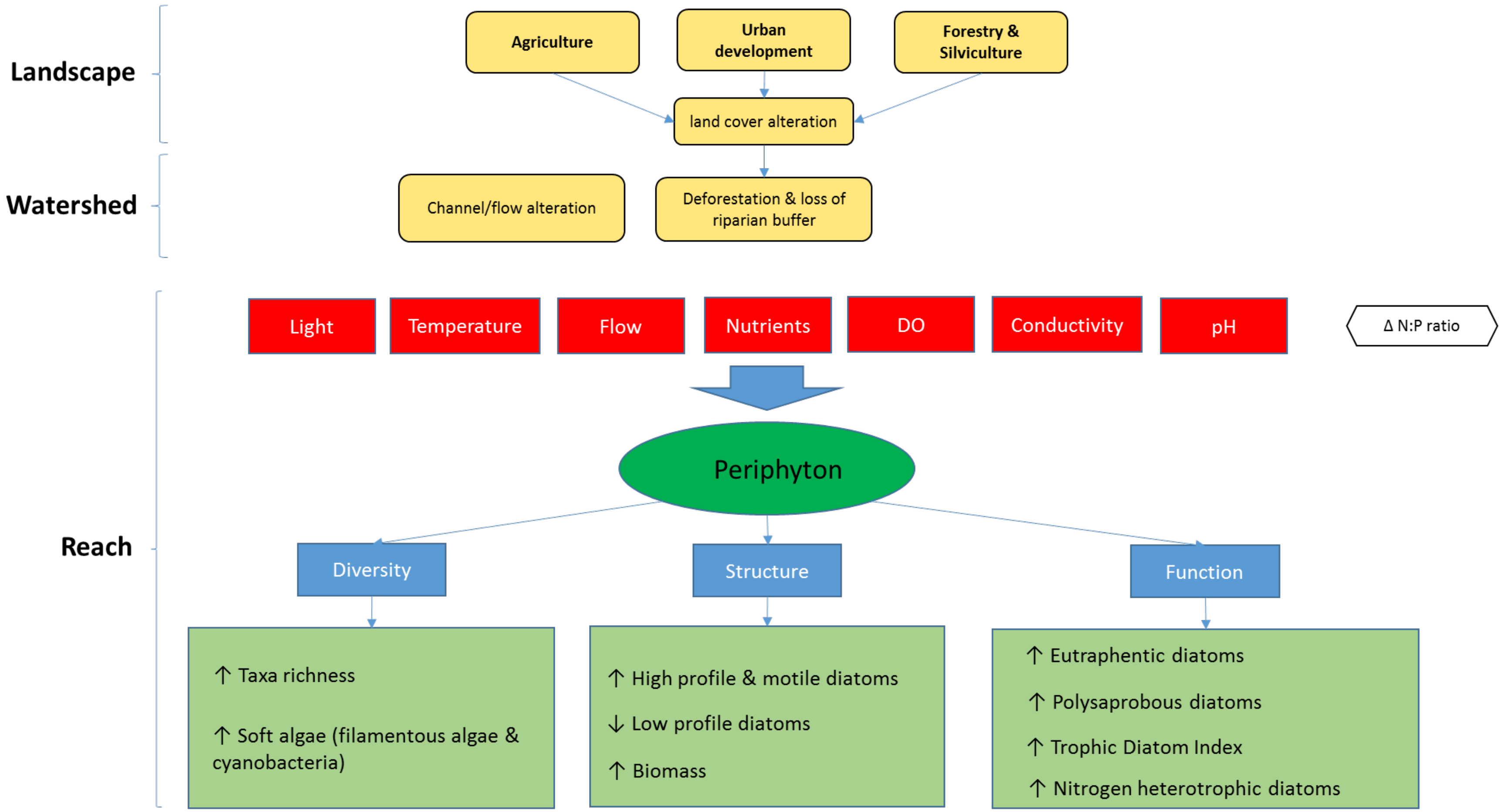
United States Environmental Protection Agency (USEPA). 2011. A primer on using biological assessments to support water quality management. U.S. Environmental Protection Agency, Office of Science and Technology, Office of Water, Washington, DC. EPA 810-R-11-01

United States Environmental Protection Agency (USEPA). 2013. Level III Ecoregions of the Conterminous United States. https://www.epa.gov/eco-research/level-iii-and-iv-ecoregions-continental-united-states

Zambrano-Bigiarini, M. 2014. hydroGOF: Goodness-of-fit functions for comparison of simulated and observed hydrological time series. R package version 0.3-8. https://CRAN.R-project.org/package=hydroGOF

Zeileis, A., T. Hothorn, and K. Hornik K. 2008. Model-Based Recursive Partitioning. Journal of Computational and Graphical Statistics 17: 492–514

APPENDIX 1. Conceptual Model



APPENDIX 2. Water quality variables used.

Table 10 -Variable names combined and method codes excluded. Abbreviations as in Table 2.

Variable name	Original variable name(s)	Method codes	Description
AFDM	AFDM, AFDM_mgm2, Biomass, periphyton, ash free dry mass, grams per square meter	All	Ash-free dry mass
Chla	Chlorophyll a, periphyton, chromatographic-fluorom; Chla; Chla_mgm2	Excluded 89	Chlorophyll a of periphyton in mg/m2
Cond	Cond_uScm, CondCal_uScm, CondHO_uScm		Specific conductivity in µS/cm
DO_mgL	DO_mg/L		Dissolved oxygen, mg/L
NH4_f		CL035, 37, 39, SHC02 (excluded 101)	Ammonium, filtered
NH4_uf	NH4_N_mgL	Excluded CL075	Ammonium, unfiltered, in mg/L as nitrogen
NO2_f	NO2_N_f_mgL	Excluded 49	Nitrite, filtered in mg/L as nitrogen
NO2_uf	NO2_N	-999, CL076, 77, 135	Nitrite, unfiltered in mg/L as nitrogen
NO3_f		All	Nitrate, filtered in mg/L as nitrogen
NO3_uf	NO3_N_mgL	-999 and ALGOR	Nitrate, unfiltered in mg/L as nitrogen
NOx_f		CLO45, 48, 50, 44, 47, 49, -999, CDR06, (excluded CL132, RED01,02)	Nitrate plus nitrite, filtered, in mg/L as nitrogen
NOx_uf	NOx_N	-999, 81, (excluded CL131)	Nitrate plus nitrite, unfiltered, in mg/L as nitrogen
OrgN_f	TKN (parameter codes 706 and 6230	Excluded CL061	Organic nitrogen, filtered in mg/L
OrgN_uf	TKN_N (parameter codes 625 and 605)		Organic nitrogen, unfiltered in mg/L
OrthoP_uf	Orthophosphate, water, unfiltered, milligrams per liter		Orthophosphate, unfiltered in mg/L
SRP	SRP_mgL; Orthophosphate, water, filtered, mg per liter		Soluble reactive phosphorus in mg/L
TempC	Temp_degC		Temperature in °C
TN_f	TN_f_mgL (parameter codes 62854 and 6020		Total nitrogen, filtered in mg/L
TN_uf	TN_N_mgL	600, 62855, 71887	Total nitrogen, unfiltered in mg/L
TP_uf	P_P_mgL, TP_P_mgL	Excluded CL84, 90, and 001	Total phosphorus, unfiltered in mg/L
TP_f	TDP_mg/L	Excluded CL52 and 60	Total phosphorus, filtered in mg/L

APPENDIX 3. Diatom metrics used.

Table 11 - List of periphyton metrics used, type (continuous or categorical) and value range, description, expected response direction to nutrient enrichment, and sources.

Field	Type, values	Description	Response to nutrients	Source
wa_OptCat_DisTotMMI	Categorical, 1 to 6	Categorical diatom indicator based on species weighted average multivariate watershed disturbance index optima.	+	1
wa_OptCat_L1DisTot	Categorical, 1 to 6	Categorical diatom indicator based on species weighted average % watershed disturbance optima.	+	1
wa_OptCat_L1Ptl	Categorical, 1 to 6	Categorical diatom indicator based on species weighted average TP optima.	+	1
wa_OptCat_LNtl	Categorical, 1 to 6	Categorical diatom indicator based on species weighted average TN optima.	+	1
wa_OptCat_LCond	Categorical, 1 to 6	Categorical diatom indicator based on species weighted average conductivity optima.	+	1
wa_OptCat_NutMMI	Categorical, 1 to 6	Categorical diatom indicator based on species weighted averaged multivariate nutrient index optima.	+	1
wa_OptCat_pH	Categorical, 1 to 6	Categorical diatom indicator based on species weighted average pH optima.	+	1
wa_OptCat_XEMBED	Categorical, 1 to 6	Categorical diatom indicator based on species weighted average % embeddedness optima.	+	1
wa_OptCat_PctFN	Categorical, 1 to 6	Categorical diatom indicator based on species weighted average % fine sediment optima.	+	1
NEWTSIC	Categorical, 1 to 6	MAIA/MAHA assessment result, Trophic State Index from oligotrophic (1) to hypertrophic (6) state	+	2
MAIATSIC	Categorical, 1 to 6	MAIA/MAHA assessment result, Trophic State Index from oligotrophic (1) to hypertrophic (6) state	+	2
Ptpv_TP_all_Ind	Categories, text	USGS NAWQA program TP indicator values for the entire dataset (1 or 2); 1 is high TP, 2 is low TP		4
Ptpv_TN_all_Ind	Categories, text	USGS NAWQA program TN indicator values for the entire dataset (1 or 2); 1 is high TN, 2 is low TN		4
Ptpv_TP_WM_Ind	Categories, text	USGS NAWQA program TP indicator values for Western Mountain (WM); 1 is high TP, 2 is low TP		4
Ptpv_TN_WM_Ind	Categories, text	USGS NAWQA program TN indicator values for Western Mountain (WM); 1 is high TN, 2 is low TN		4
Ptpv_TP_CWP_Ind	Categories, text	USGS NAWQA program TP indicator values for central, western plains (CWP); 1 is high TP, 2 is low TP		4
Ptpv_TN_CWP_Ind	Categories, text	USGS NAWQA program TN indicator values for central, western plains (CWP); 1 is high TN, 2 is low TN		4
POLL_CLASS	Categories, 1 to 3	From pollution tolerant taxa (1) to pollution sensitive taxa (4)	-	7
POLL_TOL	Categories, 1 to 5	From pollution tolerant taxa (1) to pollution sensitive taxa (5)	-	8

Field	Type, values	Description	Response to nutrients	Source
BEN_SES	Categories, 1 or 2	Sestonic (2) or benthic algae (1)	+	7
DIATAS_TP	Categories, 1 or 2	Diatom TP preference (1)		7
DIATAS_TN	Categories, 1 or 2	Diatom TN preference (1)		7

Sources:
1 = Stevenson, R. J., Y. Pan, K. M. Manoylov, C. A. Parker, D. P. Larsen, and A. T. Herlihy 2008. Development of diatom indicators of ecological conditions for streams of the western US. J. N. Am. Benthol. Soc., 2008, 27(4):1000–1016;
2 = R. J. Stevenson, Michigan State, unpublished;
3 = Stevenson, R. J. and B. Wang. 2001. Developing and Testing Algal Indicators of Nutrient Status in Florida Streams. Final Report Prepared for Florida Department of Environmental Protection. October 31, 2001.
4 = Potapova M., Charles D.F., Ponader K.C. & Winter D.M. 2004 Quantifying species indicator values for trophic diatom indices: comparison of approaches. Hydrobiologia, 517, 25 41 and Potapova M. & Charles D.F. 2007. Diatom metrics for monitoring eutrophication in rivers of the United States. Ecological Indicators, 7, 48–70;
5 = Bahls, L.L., 1993, Periphyton bioassessment methods for Montana streams: Water Quality Bureau, Department of Health and Environmental Sciences, Helena, MT, 69 p. and Porter, S.D., D.K. Mueller, N.E. Spahr, M.D. Munn. and N.M. Dubrovsky. 2008. Efficacy of algal metrics for assessing nutrient and organic enrichment in flowing waters. Freshwater Biology (2008) 53, 1036–1054;
6 = van Dam, H., Mertens, A., Sinkeldam, J., 1994. A coded and ecological indicator values of freshwater diatoms from the Netherlands. Neth. J. Aquat. Ecol. 28, 117–133.;
7 = Porter, S.D., D.K. Mueller, N.E. Spahr, M.D. Munn. and N.M. Dubrovsky. 2008. Efficacy of algal metrics for assessing nutrient and organic enrichment in flowing waters. Freshwater Biology (2008) 53, 1036–1054;
8 = Lange-Bertalot, H., 1979, Pollution tolerance of diatoms as a criterion for water quality estimation: Nova Hedwigia, v. 64, p. 285-304 and Porter, S.D., D.K. Mueller, N.E. Spahr, M.D. Munn. and N.M. Dubrovsky. 2008. Efficacy of algal metrics for assessing nutrient and organic enrichment in flowing waters. Freshwater Biology (2008) 53, 1036–1054.

APPENDIX 4. Additional diatom metrics shown in relation to nutrient concentrations.

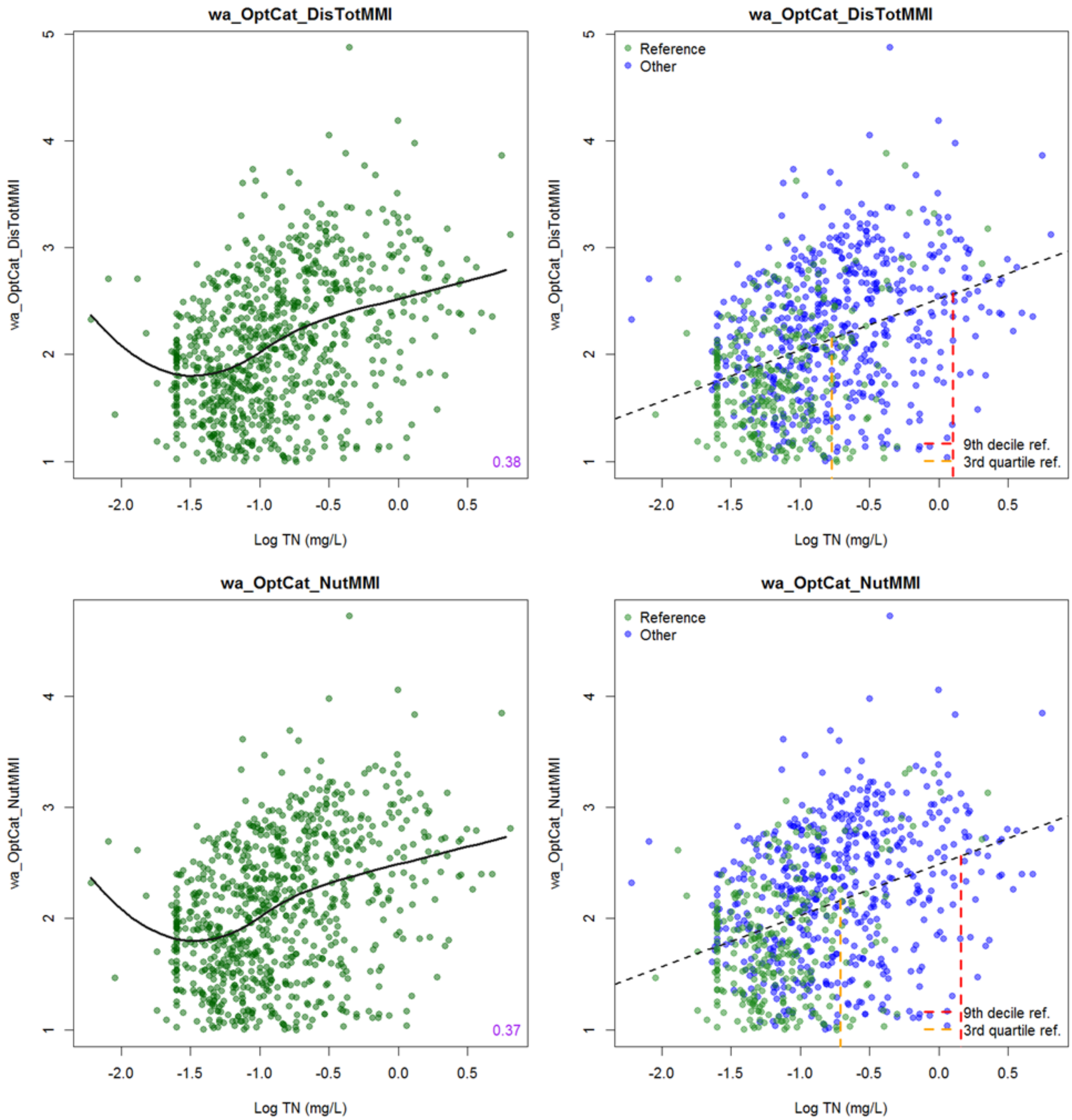


Figure 20 - Loess lines plots (left) annotated with Spearman's rho (loess span=0.75). Simple linear regression model (right) with the 9th decile and 3rd quartile of the reference site metric distribution identified and the nutrient concentrations associated with those response targets represented as the dashed lines (right). For a description of the periphyton metrics, see Methods and Appendix 2.

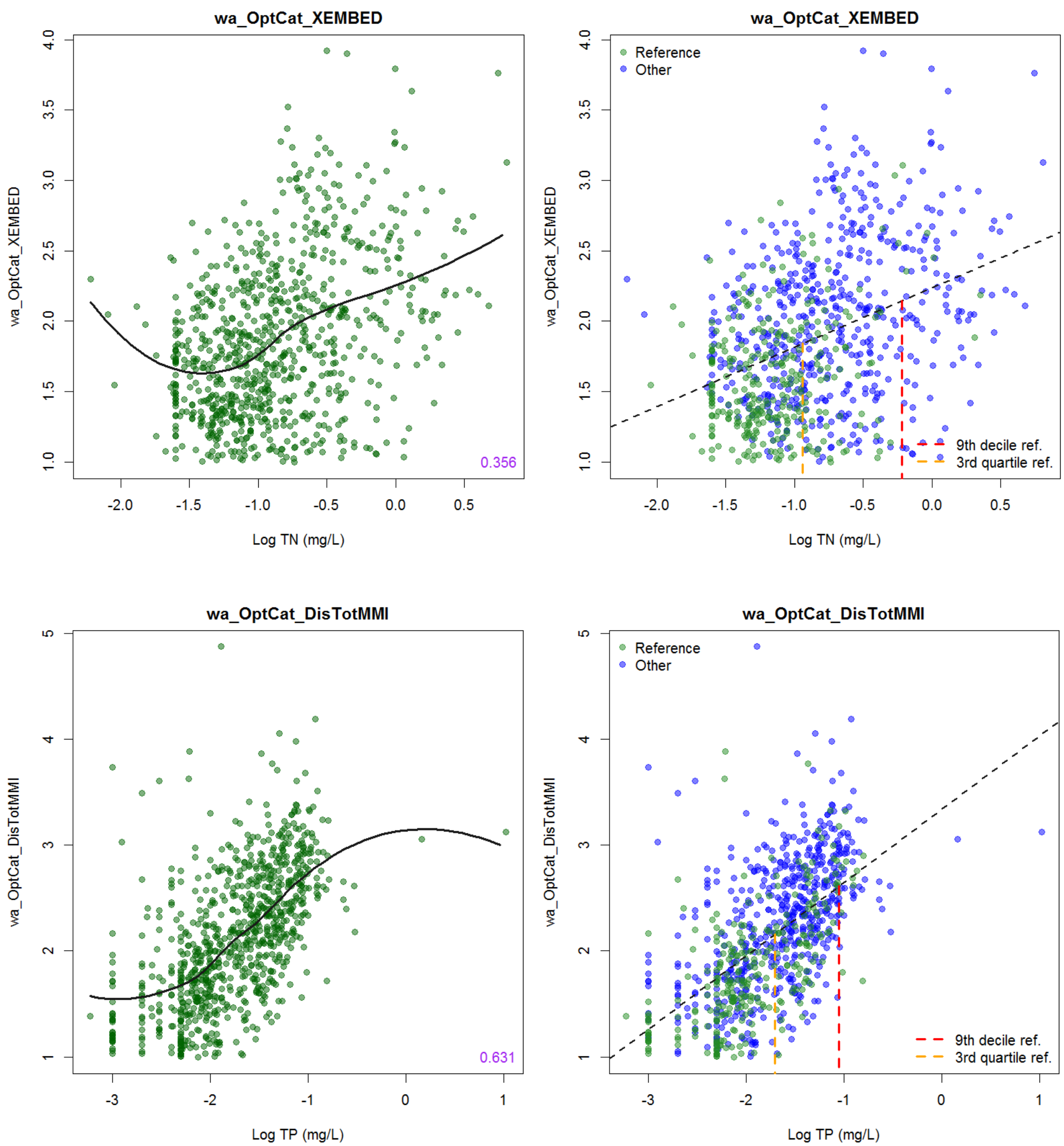


Figure 21 - Loess lines plots (left) annotated with Spearman's rho (loess span=0.75). Simple linear regression model (right) with the 9th decile and 3rd quartile of the reference site metric distribution identified and the nutrient concentrations associated with those response targets represented as the dashed lines (right). For a description of the periphyton metrics, see Methods and Appendix 2.

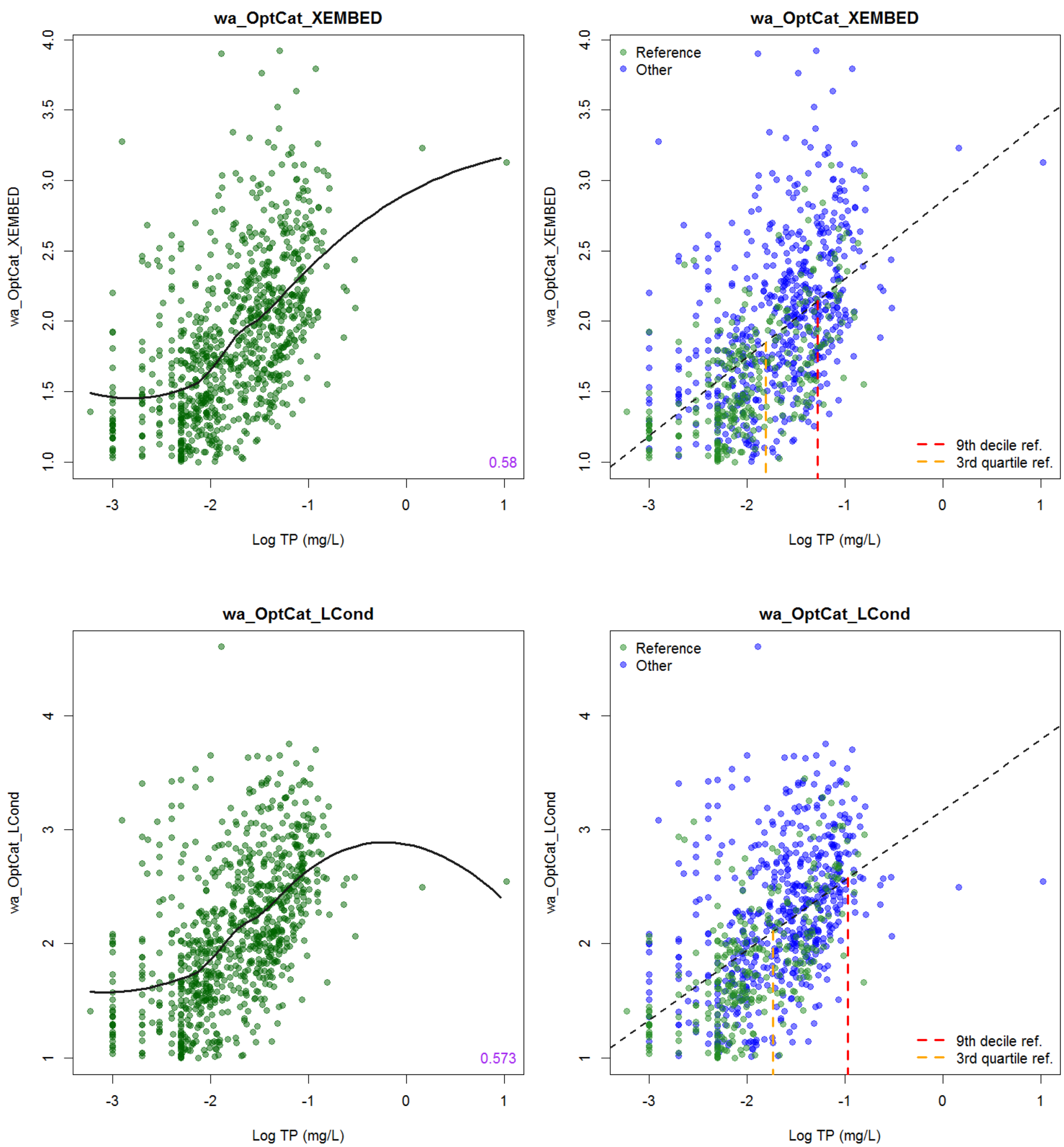


Figure 22 - Loess lines plots (left) annotated with Spearman's rho (loess span=0.75). Simple linear regression model (right) with the 9th decile and 3rd quartile of the reference site metric distribution identified and the nutrient concentrations associated with those response targets represented as the dashed lines (right). For a description of the periphyton metrics, see Methods and Appendix 2.

APPENDIX 5. Classification analysis detail.

Residual analysis

For classification, we looked at three independent, supporting lines of analysis: residual analysis, recursive partitioning, and Nash-Sutcliffe efficiency. Residual analysis revealed that diatom metric - nutrient concentration relationships varied along gradients of latitude and precipitation for TN-responsive metrics (e.g., Figure 23 and Figure 24). Variability in residuals was less along longitude and elevation gradients and did not differ from zero. For TP responsive metrics, variability in residuals was greatest along the precipitation gradient (Figure 26 and Figure 27). Ecoregions 3 (Willamette Valley), 10 (Columbia Plateau), and 11 (Blue Mountains) tended to have higher biased residuals than other ecoregions for both nutrients, but even more so for TN (Figure 25 and Figure 28).

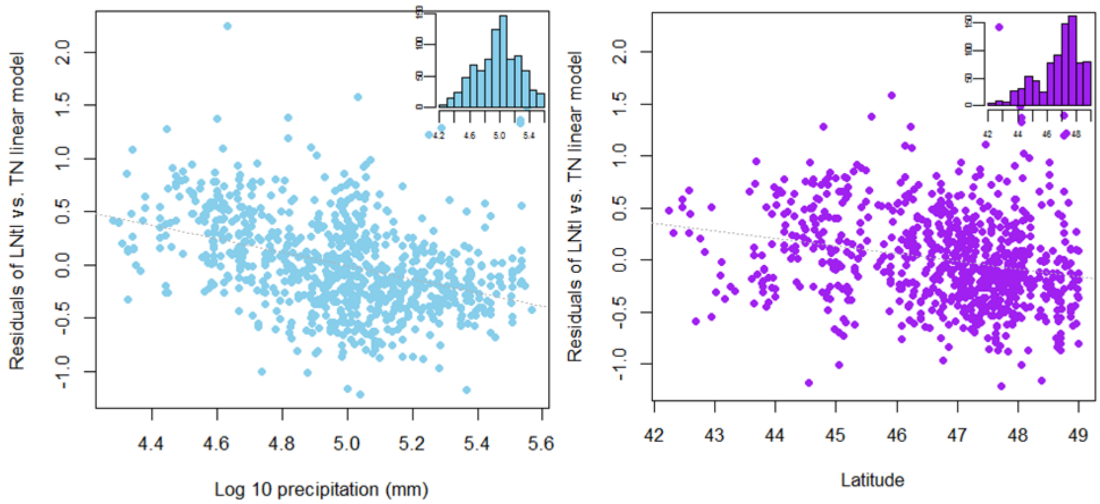


Figure 23 - Residuals of the weighted average LNtl diatom metric (wa_OptCat_LNtl) vs. log10 transformed TN linear model as a function of PRISM precipitation (left) and latitude (right).

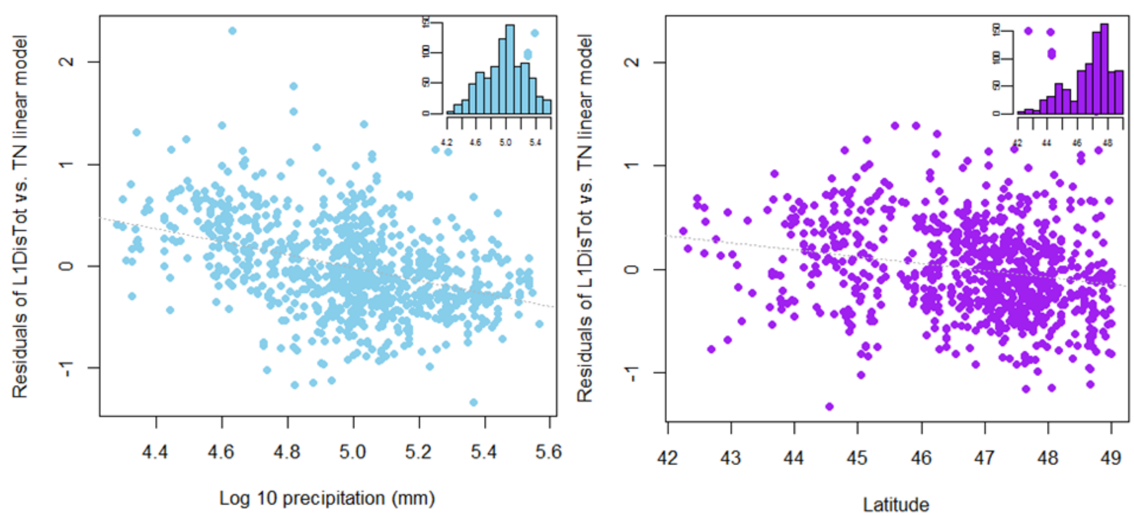


Figure 24 - Residuals of the weighted average L1DisTot diatom metric (wa_OptCat_L1DisTot) vs. log$_{10}$ transformed TN linear model as a function of PRISM precipitation (top left), elevation (top right), latitude and longitude (bottom left and right, respectively).

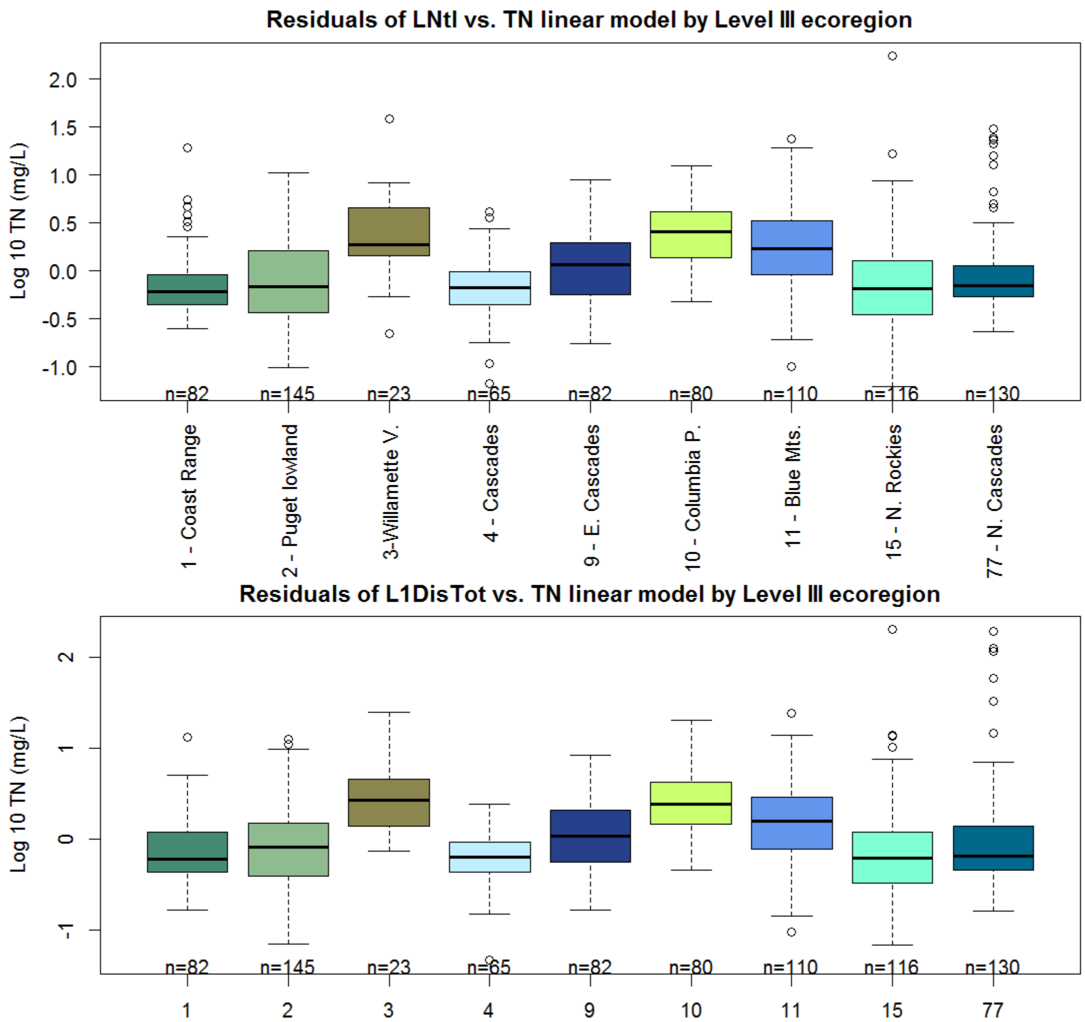


Figure 25 - Residuals of the weighted average LNtl diatom metric (wa_OptCat_LNtl) vs. $\log_{10}$ transformed TN linear model (top) and weighted average L1DisTot diatom metric (wa_OptCat_L1DisTot) vs. $\log_{10}$ transformed TN model (bottom) as a function of Omernik Level III ecoregion.

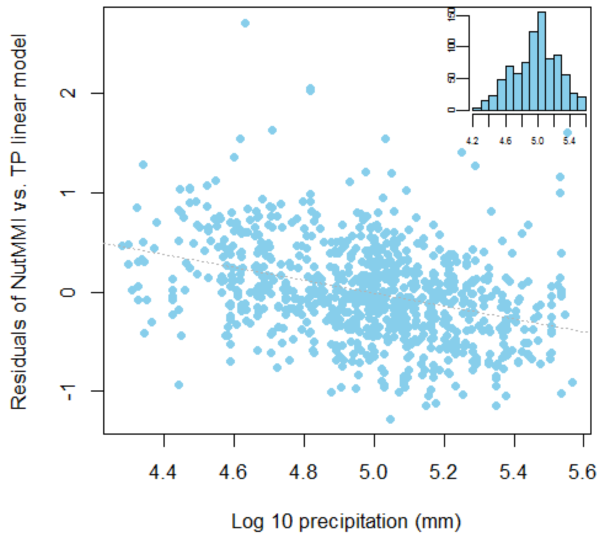


Figure 26 - Residuals of the weighted average NutMMI diatom metric (wa_OptCat_NutMMI) vs. $\log_{10}$ transformed TN linear model as a function of PRISM precipitation (top left), elevation (top right), latitude and longitude (bottom left and right, respectively).

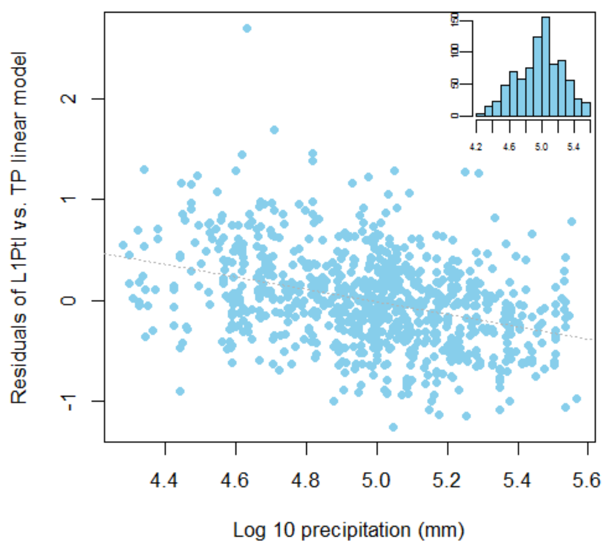


Figure 27 - Residuals of the weighted average L1Ptl diatom metric (wa_OptCat_L1Ptl) vs. log10 transformed TN linear model as a function of PRISM precipitation (top left), elevation (top right), latitude and longitude (bottom left and right, respectively).

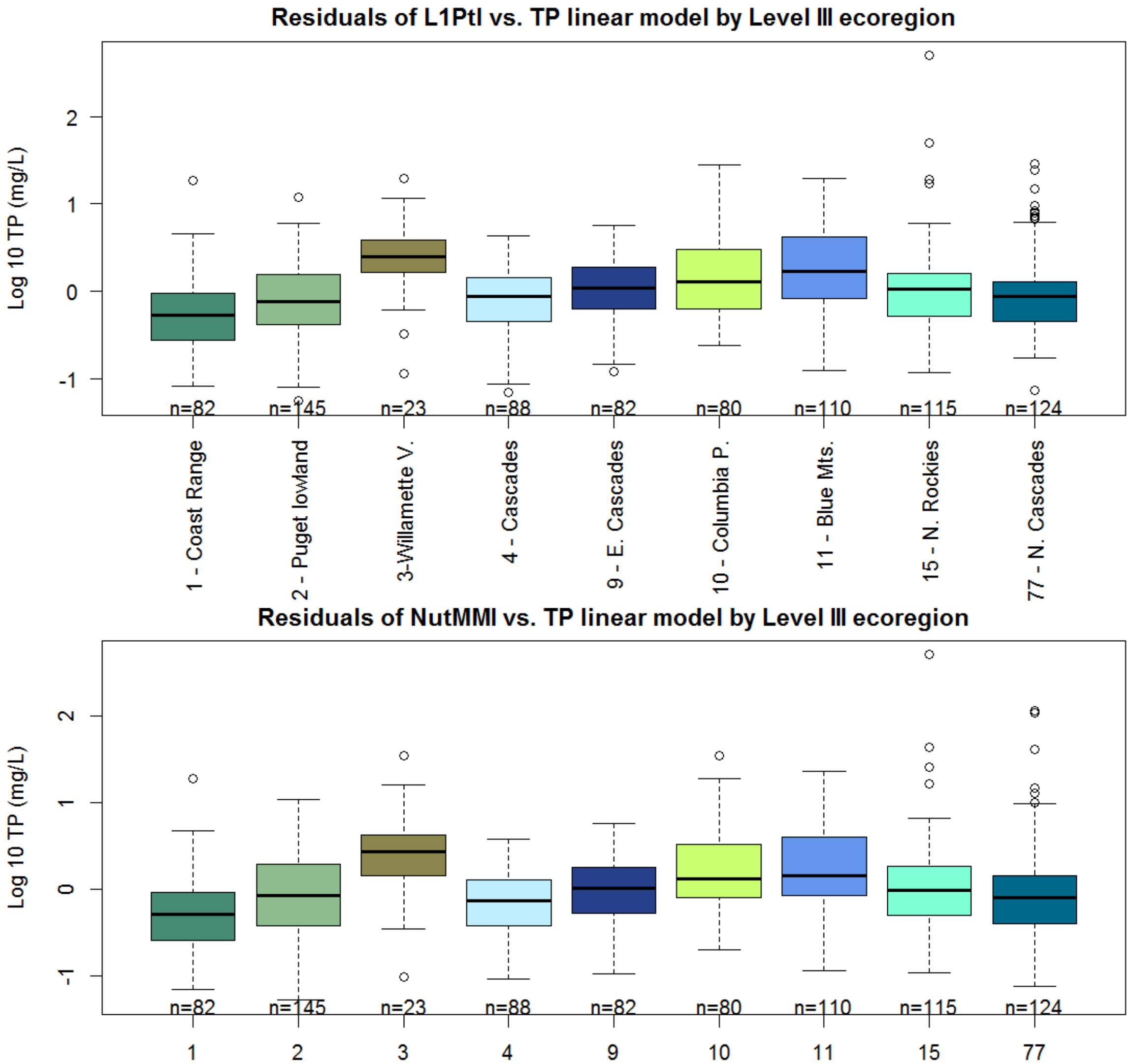


Figure 28 - Residuals of the weighted average L1Ptl diatom metric (wa_OptCat_L1Ptl) vs. $\log_{10}$ transformed TP linear model (top) and weighted average NutMMI diatom metric (wa_OptCat_NutMMI) vs. $\log_{10}$ transformed TP model (bottom) as a function of Omernik Level III ecoregion.

Model-based recursive partitioning

Model based-recursive partitioning for metrics responsive to TN indicated that latitude splits around 46.4° (South of Yakima) produced different models, with steeper slopes characteristic of lower latitudes (Figure 29). Model based-recursive partitioning for metrics responsive to TP produced latitude splits slightly higher, around 47.4° (between Tacoma and Seattle) (Figure 30). Splits in longitude for both nutrient models occurred between -121.2° to -121.8°. When ecoregion was used as the splitting variable, models for both nutrients spilt with ecoregions 3, 10 and 11 (Willamette Valley, Columbia Plateau, and Blue Mountains, respectively) apart from the rest (Figure 31 and Figure 32). Additional model output shown in Figure 33 to Figure 37.

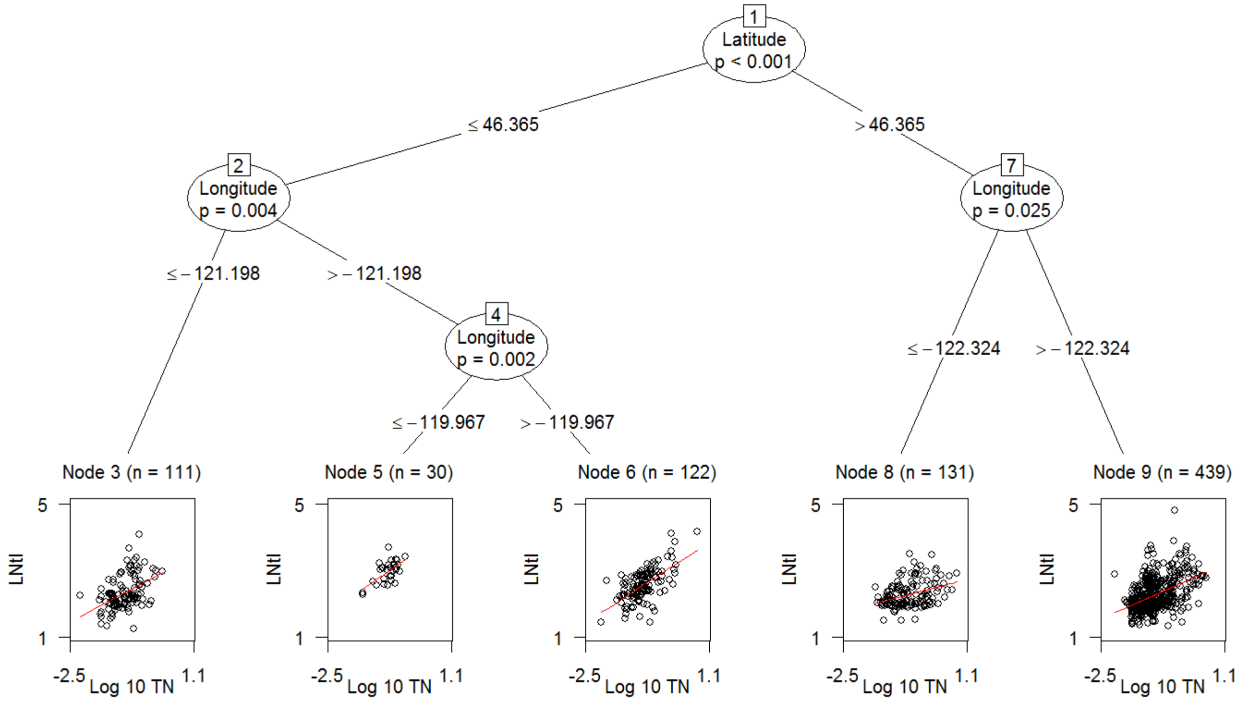


Figure 29 - Model-based recursive partitioning of the wa_OptCat_LNtl diatom optima metric as a response to TN concentration using longitude and latitude as potential splitting variables.

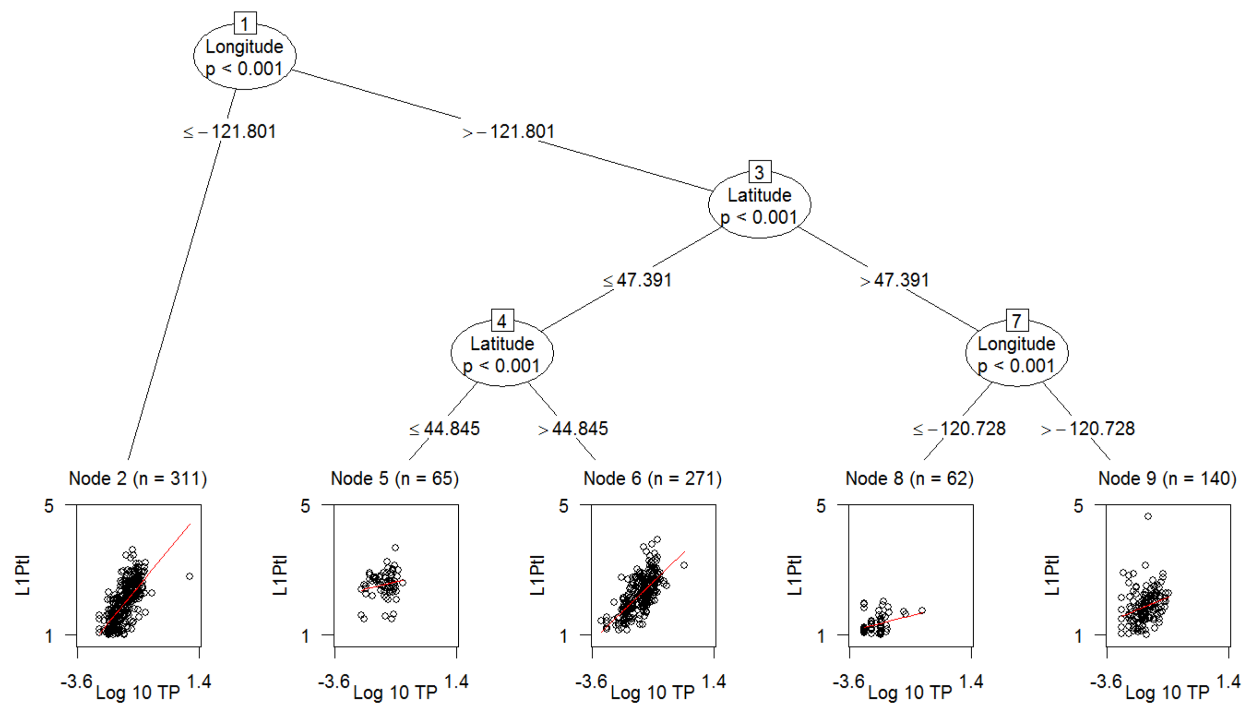


Figure 30 - Model-based recursive partitioning of the wa_OptCat_L1Ptl diatom optima metric as a response to TP concentration using longitude and latitude as potential splitting variables.

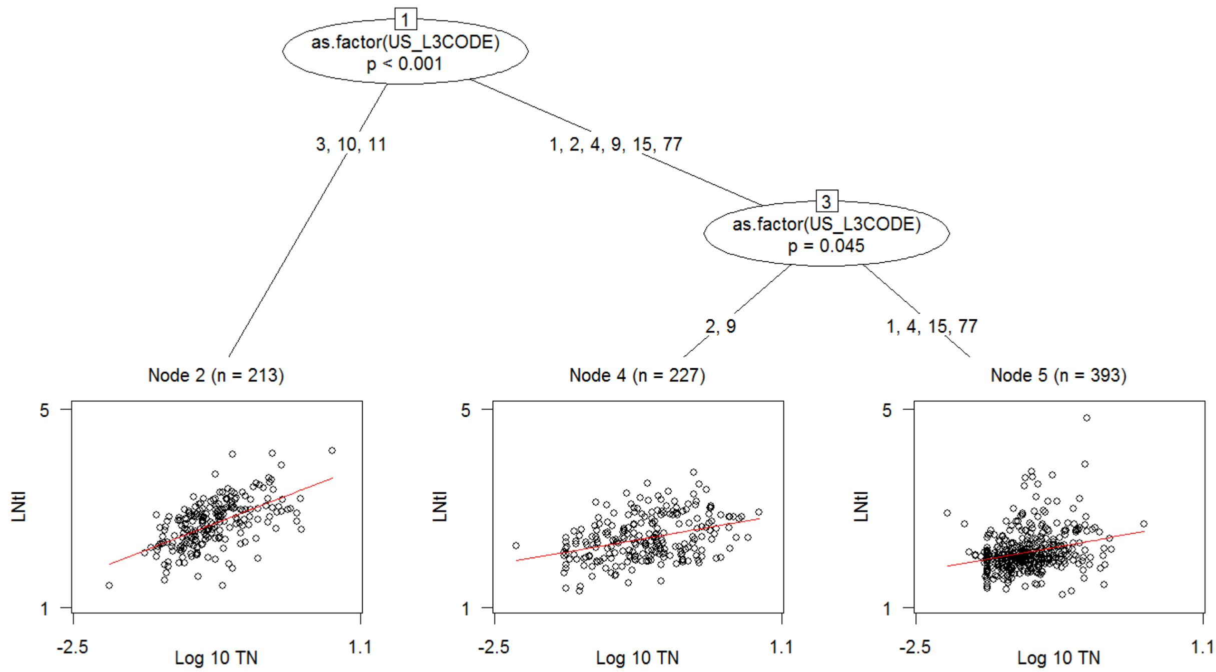


Figure 31 - Model-based recursive partitioning of the wa_OptCat_LNtl optima metric as a response to TN concentration using Omernik Level III ecoregion as the splitting variable.

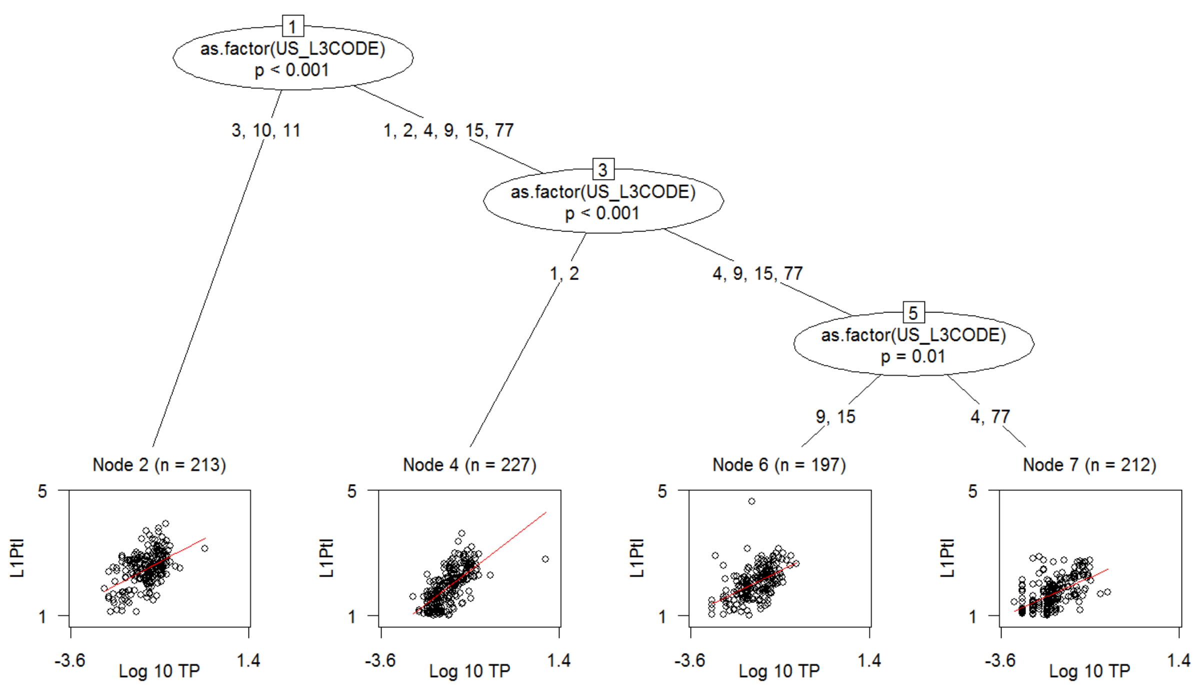


Figure 32 - Model-based recursive partitioning of the wa_OptCat_L1Ptl optima metric as a response to TP concentration using Omernik Level III ecoregion as the splitting variable.

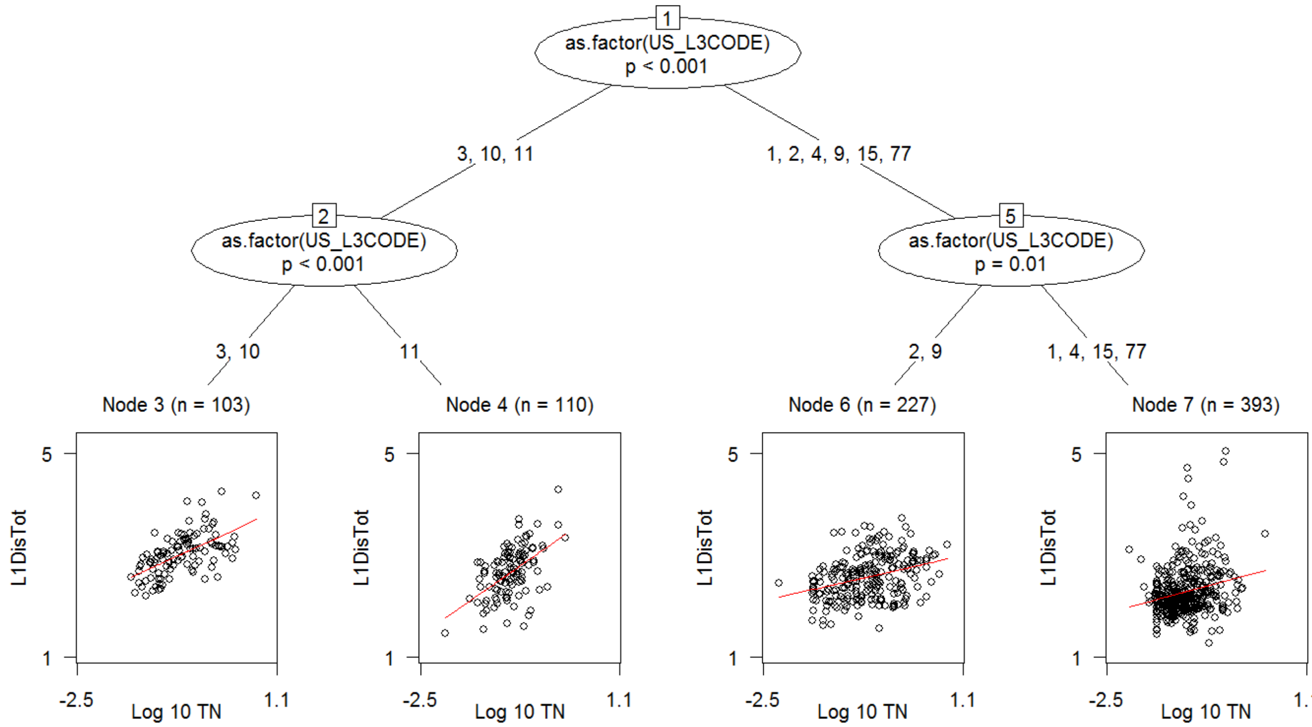


Figure 33 - Model-based recursive partitioning of the Western EMAP diatom optima metric to watershed disturbance (OptCat_L1DisTot) as a response to TN concentration using Omernik Level III ecoregion as the splitting variable.

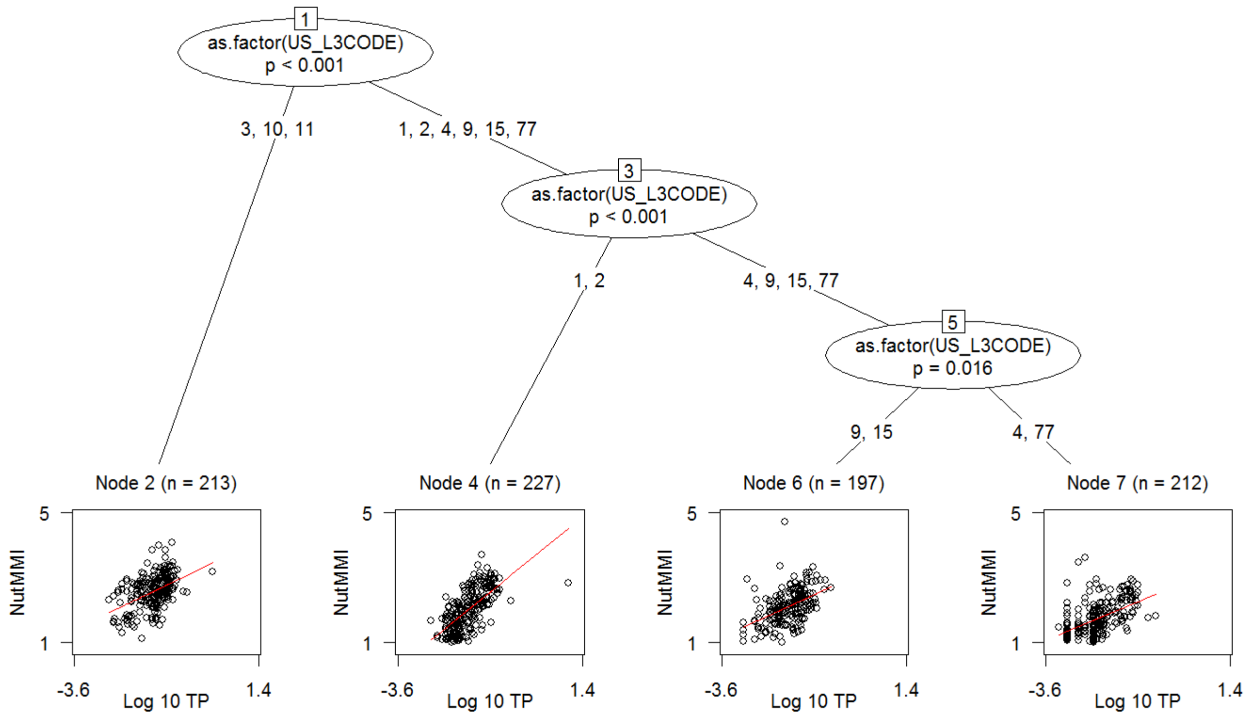


Figure 34 - Model-based recursive partitioning of the Western EMAP diatom MMI periphyton metric (OptCat_NutMMI) as a response to TP concentration using Omernik Level III ecoregion as the splitting variable.

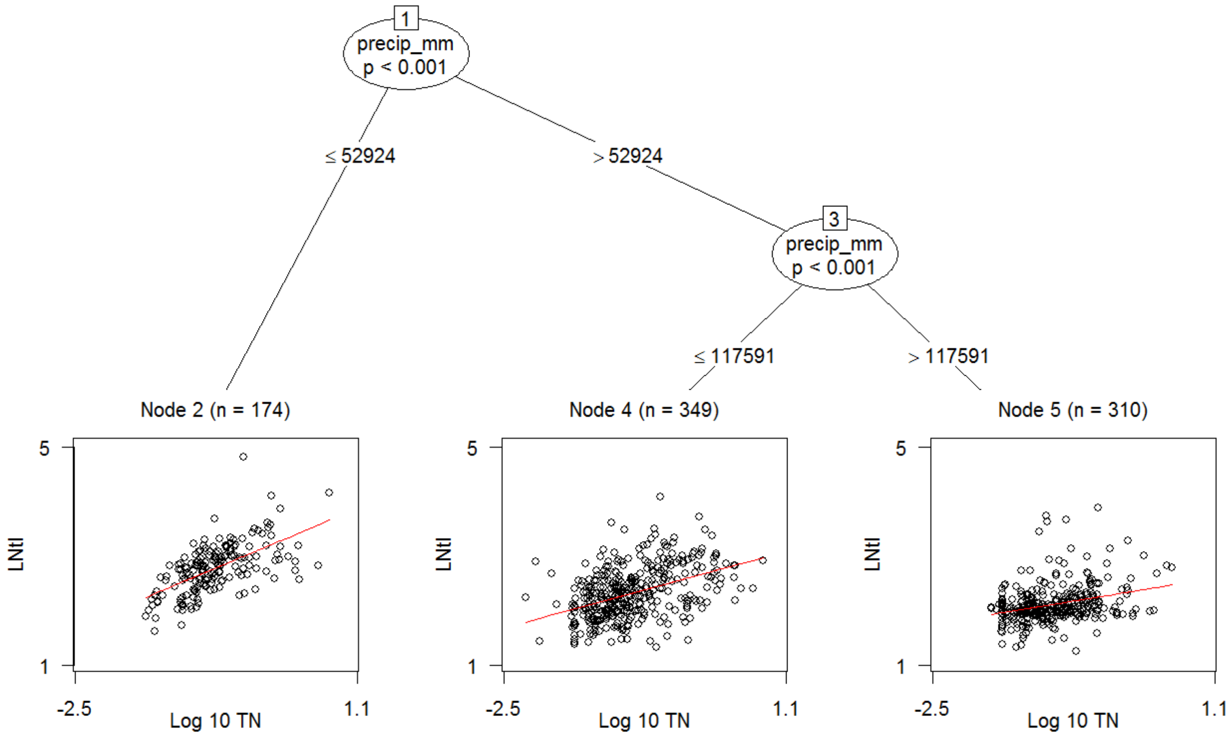


Figure 35 - Model-based recursive partitioning of the Western EMAP diatom optima metric to log TN (OptCat_LNtl) as a response to TN concentration using precipitation as the splitting variable.

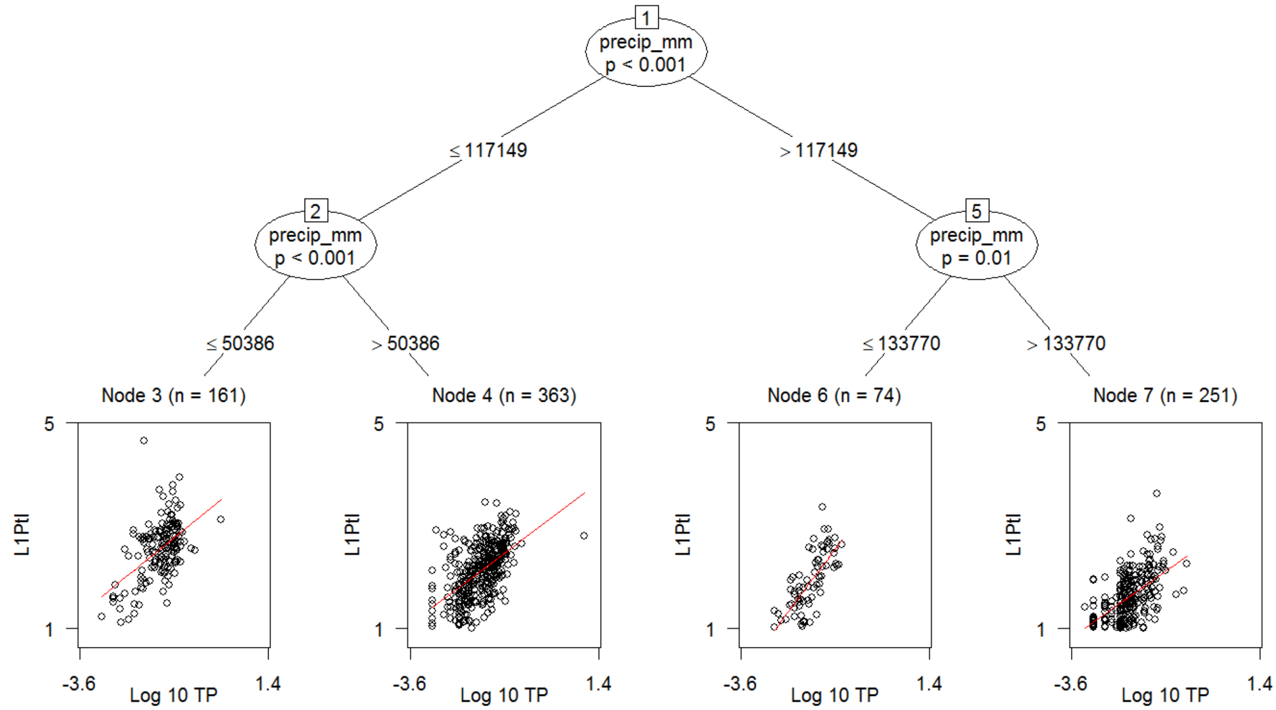


Figure 36 - Model-based recursive partitioning of the Western EMAP diatom optima metric to log TP (OptCat_L1Ptl) as a response to TP concentration using precipitation (units are hundredths of mm) as the splitting variable.

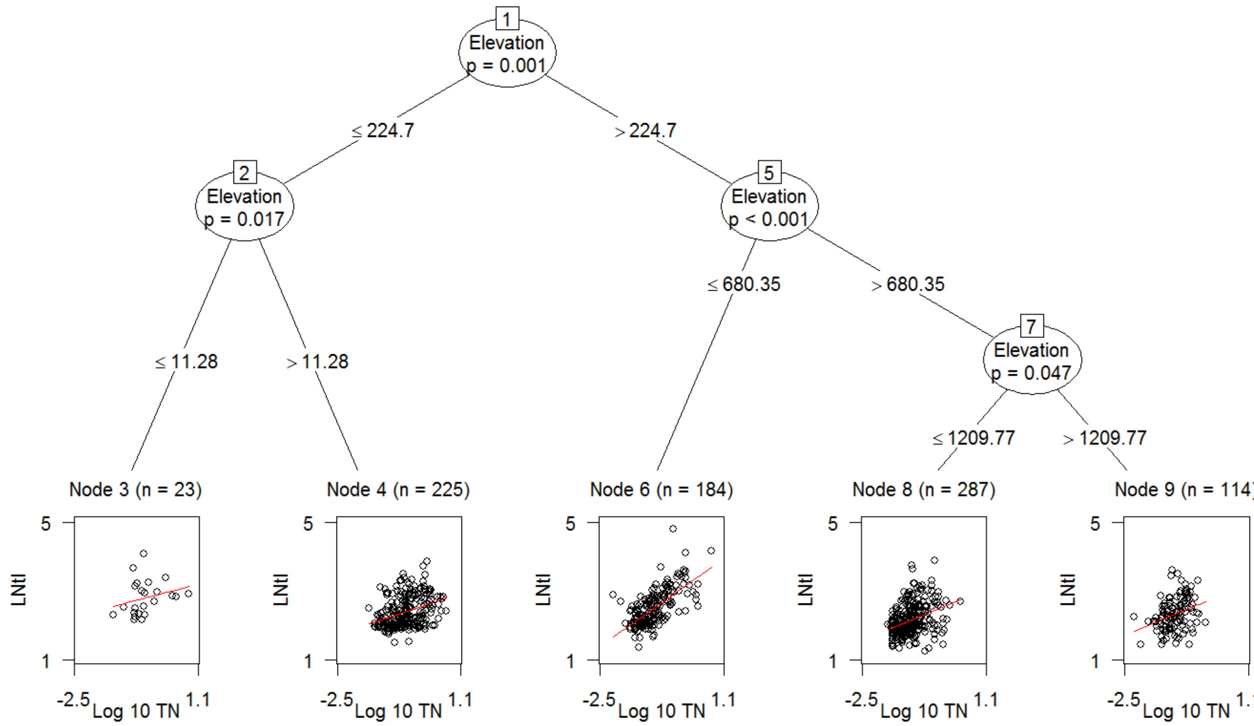


Figure 37 - Model-based recursive partitioning of the Western EMAP diatom optima metric to log TN (OptCat_LNtl) as a response to TN concentration using elevation as the splitting variable.

Nash-Sutcliffe efficiency calculation

The NSE for models built from metrics sensitive to TN and log TN concentration using ecoregions 3, 10, and 11 (Willamette Valley, Columbia Plateau, and Blue Mountains, respectively) had values around 0.34. The NSE values for the remaining data when compared to the ecoregion 3, 10 and 11 models were negative, indicating that the observed (test) data mean was a better estimate than the ecoregion 3, 10, and 11 model (Figure 38, top). The difference in response may be due to greater P supply in the latter regions or other factors influencing N sensitivity.

Models built from metrics sensitive to log TP concentration showed a different pattern; the ecoregion 3, 10 and 11 models and lower NSE's (0.245 for OptCat_L1Ptl), and the remaining ecoregions had negative NSE's however, based on linear model fits, the TP models appeared to differ in intercept rather than slope (Figure 38, bottom).

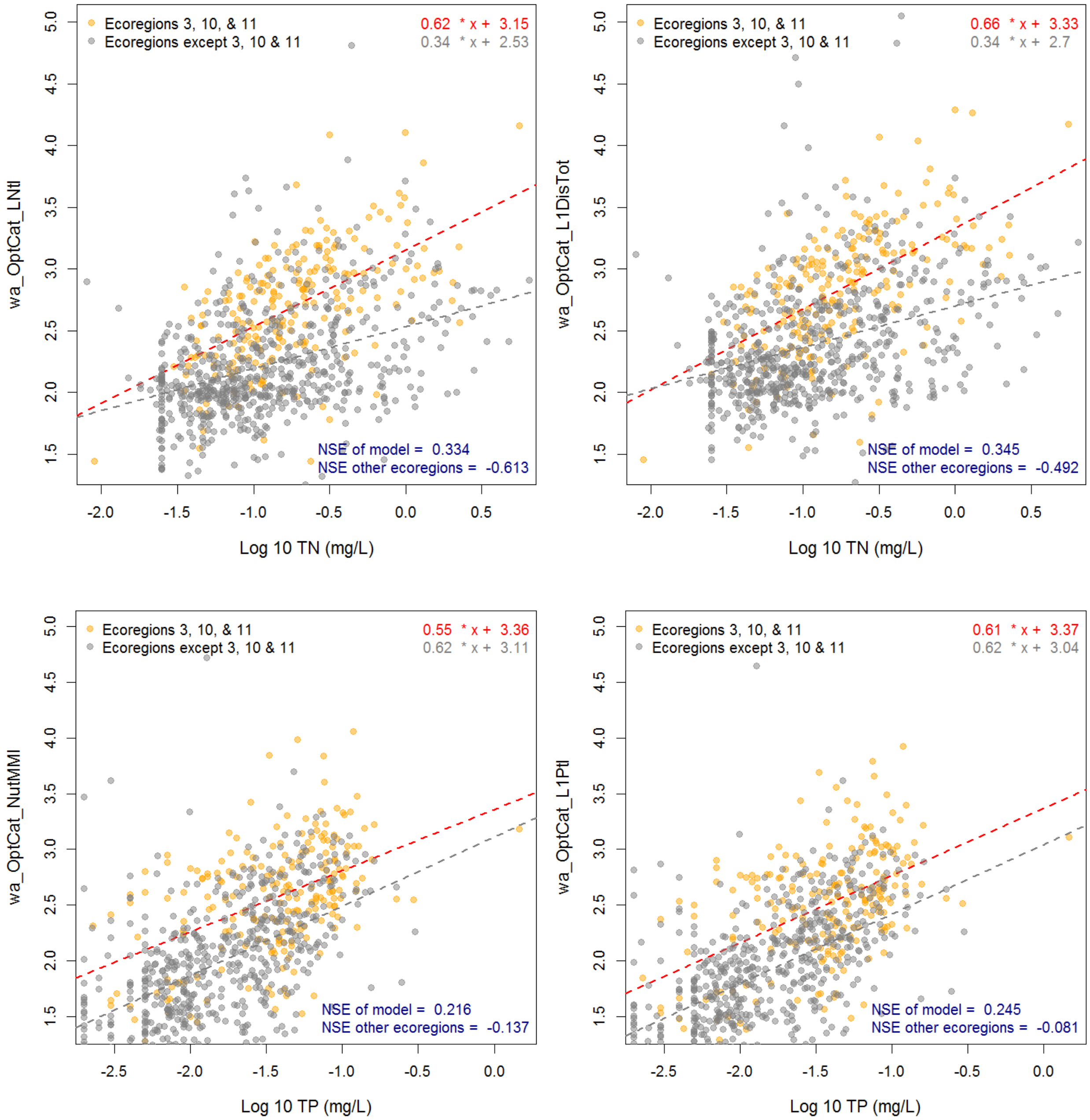


Figure 38 - Diatom metrics from sampling stations from ecoregions 3, 10, and 11 (Willamette Valley, Columbia Plateau, and Blue Mountains, respectively) were used to build models against which models built from samples in all remaining ecoregions were tested. Top performing TN metrics (top) and TP metrics (bottom).